



TUGAS AKHIR - SS141501

**FAKTOR-FAKTOR YANG MEMPENGARUHI
ANGKA KECELAKAAN LALU LINTAS
DI PROVINSI JAWA TIMUR MENGGUNAKAN
REGRESI NONPARAMETRIK SPLINE *TRUNCATED***

**NYM CISTA STRIRATNA DEWI
NRP 06211645000011**

**Dosen Pembimbing
Prof. Dr. Drs. I Nyoman Budiantara, M.Si.**

**PROGRAM STUDI SARJANA
DEPARTEMEN STATISTIKA
FAKULTAS MATEMATIKA, KOMPUTASI, DAN SAINS DATA
INSTITUT TEKNOLOGI SEPULUH NOPEMBER
SURABAYA 2018**



TUGAS AKHIR - SS141501

**FAKTOR-FAKTOR YANG MEMPENGARUHI
ANGKA KECELAKAAN LALU LINTAS
DI PROVINSI JAWA TIMUR MENGGUNAKAN
REGRESI NONPARAMETRIK SPLINE *TRUNCATED***

**NYM CISTA STRIRATNA DEWI
NRP 06211645000011**

**Dosen Pembimbing
Prof. Dr. Drs. I Nyoman Budiantara, M.Si.**

**PROGRAM STUDI SARJANA
DEPARTEMEN STATISTIKA
FAKULTAS MATEMATIKA, KOMPUTASI, DAN SAINS DATA
INSTITUT TEKNOLOGI SEPULUH NOPEMBER
SURABAYA 2018**



FINAL PROJECT - SS141501

**FACTORS THAT INFLUENCE TRAFFIC ACCIDENTS
IN EAST JAVA USING NONPARAMETRIC SPLINE
TRUNCATED REGRESSION**

**NYM CISTA STRIRATNA DEWI
SN 06211645000011**

**Supervisor:
Prof. Dr. Drs. I Nyoman Budiantara, M.Si.**

**UNDERGRADUATE PROGRAMME
DEPARTMENT OF STATISTICS
FACULTY OF MATHEMATICS, COMPUTING, AND DATA SCIENCE
INSTITUT TEKNOLOGI SEPULUH NOPEMBER
SURABAYA 2018**

LEMBAR PENGESAHAN

FAKTOR-FAKTOR YANG MEMPENGARUHI ANGKA KECELAKAAN LALU LINTAS DI PROVINSI JAWA TIMUR MENGGUNAKAN REGRESI NONPARAMETRIK SPLINE *TRUNCATED*

TUGAS AKHIR

Diajukan Untuk Memenuhi Salah Satu Syarat
Memperoleh Gelar Sarjana Sains
pada

Program Studi Sarjana Departemen Statistika
Fakultas Matematika, Komputasi, dan Sains Data
Institut Teknologi Sepuluh Nopember

Oleh :

Nym Cista Striratna Dewi
NRP. 062116 4500 0011

Disetujui oleh Pembimbing:

Prof. Dr. Drs. I Nyoman Budiantara, M.Si.
NIP. 19650603 198903 1 003



Mengetahui,
Kepala Departemen



Dr. Suhartono
NIP. 19710929 199512 1 001

SURABAYA, JULI 2018

**FAKTOR-FAKTOR YANG MEMPENGARUHI
ANGKA KECELAKAAN LALU LINTAS
DI PROVINSI JAWA TIMUR MENGGUNAKAN
REGRESI NONPARAMETRIK SPLINE TRUNCATED**

Nama Mahasiswa : Nym Cista Striratna Dewi
NRP : 06211645000011
Departemen : Statistika-FMKSD-ITS
Dosen Pembimbing : Prof. Dr. Drs. I Nyoman Budiantara, M.Si.

Abstrak

Jawa Timur merupakan salah satu provinsi yang memiliki kepadatan penduduk yang tinggi di Indonesia. Kepadatan penduduk yang tinggi memicu adanya permasalahan yaitu kecelakaan lalu lintas. Selama tiga tahun berturut-turut yakni pada tahun 2012 hingga 2014, Jawa Timur menempati urutan pertama dengan kasus kecelakaan lalu lintas paling tinggi di Indonesia. Tingginya kasus kecelakaan lalu lintas yang terjadi di Jawa Timur disebabkan oleh banyak faktor. Oleh sebab itu, dilakukan penelitian untuk mengetahui faktor penyebab terjadinya kecelakaan lalu lintas di Jawa Timur. Pada penelitian ini, angka kecelakaan lalu lintas dan faktor-faktor yang mempengaruhinya tidak membentuk suatu pola tertentu, sehingga digunakan metode regresi nonparametrik Spline Truncated. Model terbaik didapatkan dari titik knot optimum berdasarkan nilai Generalized Cross Validation (GCV) minimum. Model regresi nonparametrik Spline terbaik adalah dengan menggunakan kombinasi titik knot (1,3,3,3,3). Kelima variabel prediktor yang digunakan berpengaruh signifikan terhadap angka kecelakaan lalu lintas di Jawa Timur, yaitu kepadatan penduduk, persentase pelaku usia remaja, persentase pendidikan terakhir pelaku lebih dari SMP, rasio jenis kelamin pelaku, dan persentase pelaku kecelakaan tidak memiliki SIM. Nilai koefisien determinasi yang dihasilkan dari model ini sebesar 86,73%.

Kata kunci: *Angka Kecelakaan Lalu Lintas, GCV, Regresi Nonparametrik, Spline Truncated, Titik Knot*

(Halaman ini sengaja dikosongkan)

FACTORS THAT INFLUENCE TRAFFIC ACCIDENTS IN EAST JAVA USING NONPARAMETRIC SPLINE TRUNCATED REGRESSION

Name : Nym Cista Striratna Dewi
Student Number : 06211645000011
Department : Statistics
Supervisor : Prof. Dr. Drs. I Nyoman Budiantara, M.Si

Abstract

East Java is one of the province with high population density in Indonesia. This high population density triggered the problem of traffic accidents. For three consecutive years, East Java became the province with the highest number of traffic accidents in Indonesia. The high number of traffic accidents in East Java is caused by many factors. Therefore, research will be conducted to determine the factors causing traffic accidents in East Java. In this research, the number of traffic accidents and the factors that influence does not form a certain pattern, so that, the Spline nonparametric regression method is used. The best model is obtained from the optimum knot point based on the minimum value of Generalized Cross Validation (GCV). The best Spline nonparametric regression model is using a combination of knots (1,3,3,3,3). All predictor variables significantly influence the model, that is population density, percentage of teenage offender, percentage of last education of perpetrator more than junior high, sex ratio of perpetrator, and percentage of perpetrator who does not have a driver's license. Determination coefficient value from this model equal to 86,73%.

Keywords: *GCV, Knot Points, Nonparametric Regression, Spline Truncated, Traffic Accidents*

(This page intentionally left blank)

KATA PENGANTAR

Puji syukur Astungkara, penulis panjatkan kepada Ida Sang Hyang Widhi Wasa, Tuhan Yang Maha Esa atas limpahan rahmat-Nya sehingga penulis dapat menyelesaikan Tugas Akhir yang berjudul “Faktor-Faktor Yang Mempengaruhi Angka Kecelakaan Lalu Lintas Di Provinsi Jawa Timur Menggunakan Regresi Nonparametrik Spline *Truncated*” dengan lancar.

Penulis menyadari bahwa dalam penyusunan Tugas Akhir ini tidak terlepas dari bantuan dan dukungan dari berbagai pihak. Oleh sebab itu, penulis mengucapkan terima kasih yang sebesar-besarnya kepada :

1. Dr. Suhartono selaku Kepala Departemen Statistika FMKSD ITS.
2. Dr. Sutikno, S.Si, M.Si. selaku Kepala Program Studi S1 Departemen Statistika FMKSD ITS.
3. Prof. Dr. Drs. I Nyoman Budiantara, M.Si. selaku dosen pembimbing yang telah meluangkan waktunya kepada penulis untuk memberikan ilmu dan pengarahan dengan sabar.
4. Dr. Ismaini Zain, M.Si. dan Dra. Madu Ratna, M.Si. selaku dosen penguji yang telah memberikan kritik dan saran untuk kesempurnaan tugas akhir ini.
5. Dr. Kartika Fithriarsari, M.Si. selaku dosen wali yang telah meluangkan waktu untuk membimbing penulis selama masa perkuliahan.
6. Seluruh bapak dan ibu dosen Statistika ITS yang telah memberikan ilmu, wawasan, serta bimbingan yang tidak terbatas, serta segenap seluruh karyawan Departemen Statistika ITS atas bantuan terkait akademik.
7. Kepolisian Daerah Jawa Timur (Polda Jatim) khususnya divisi Laka Lantas, atas segala informasi dan bantuan dalam pemenuhan data tugas akhir.
8. Kedua orang tua dan saudara kandung penulis (Mitha, Winda, & Putri) yang selalu memberikan doa, kasih sayang, dan semangat selama ini.

9. Teman seperjuangan Spline: Ani, Sabella, Camelia, dan Beti atas diskusi tugas akhir selama satu semester.
10. Partner terbaik selama masa perkuliahan: Sania, Nicea, Putri, Yongky, dan Ijah.
11. Teman Selusin (Kelas A), atas kekompakan dan lelucon yang tidak ada habisnya.
12. Seluruh keluarga besar LJ Statistika ITS angkatan 2016, untuk kebersamaannya selama ini.
13. Motivator penulis, Kak Angga, atas segala saran, nasehat serta dorongan untuk segera menyelesaikan tugas akhir ini.
14. Serta pihak-pihak lain yang tidak dapat penulis sebutkan satu persatu.

Penulis menyadari masih banyak kekurangan dalam laporan ini, maka saran dan kritik dari pembaca demi perbaikan sangat dibutuhkan oleh penulis. Semoga tugas akhir ini dapat bermanfaat bagi pembaca dan pihak lainnya.

Surabaya, Juli 2018

Penulis

DAFTAR ISI

	Halaman
HALAMAN JUDUL	i
<i>TITLE PAGE</i>	ii
LEMBAR PENGESAHAN	iii
ABSTRAK	v
<i>ABSTRACT</i>	vii
KATA PENGANTAR	ix
DAFTAR ISI	xi
DAFTAR TABEL	xiii
DAFTAR GAMBAR	xv
DAFTAR LAMPIRAN	xvii
 BAB I PENDAHULUAN	
1.1 Latar Belakang	1
1.2 Rumusan Masalah	4
1.3 Tujuan Penelitian	5
1.4 Manfaat Penelitian	5
1.5 Batasan Masalah	5
 BAB II TINJAUAN PUSTAKA	
2.1 Statistika Deskriptif	7
2.2 Analisis Regresi	7
2.3 Regresi Nonparametrik Spline	8
2.4 Pemilihan Titik Knot Optimal	11
2.5 Pengujian Parameter Model	11
2.6.1 Pengujian Secara Serentak	11
2.6.2 Pengujian Secara Parsial	12
2.6 Pengujian Asumsi Residual	13
2.8.1 Uji Asumsi Identik	13
2.8.2 Uji Asumsi Independen	14
2.8.3 Uji Asumsi Berdistribusi Normal	15
2.7 Kriteria Pemilihan Model Terbaik	16
2.8 Angka Kecelakaan Lalu Lintas	16
2.9 Faktor Penyebab Kecelakaan Lalu Lintas	17

BAB III METODOLOGI PENELITIAN

3.1 Sumber Data.....	21
3.2 Variabel Penelitian	21
3.3 Definisi Operasional.....	22
3.4 Struktur Data	23
3.5 Langkah Analisis.....	23
3.6 Diagram Alir	24

BAB IV ANALISIS DAN PEMBAHASAN

4.1 Karakteristik Angka Kecelakaan Lalu Lintas di Provinsi Jawa Timur.....	27
4.1.1 Analisis Deskriptif Angka Kecelakaan Lalu Lintas di Provinsi Jawa Timur	27
4.1.2 Analisis Pola Hubungan Angka Kecelakaan Lalu Lintas dengan Faktor-Faktor yang Diduga Mempengaruhi	33
4.2 Pemodelan Angka Kecelakaan Lalu Lintas di Jawa Timur Menggunakan Regresi Nonparametrik Spline.....	36
4.2.1 Memodelkan Menggunakan Titik Knot	37
4.2.2 Pemilihan Titik Knot Optimal.....	43
4.2.3 Model Terbaik untuk Angka Kecelakaan Lalu Lintas di Provinsi Jawa Timur	44
4.2.4 Pengujian Signifikansi Parameter Model Regresi Nonparametrik Spline	44
4.2.5 Pengujian Asumsi Residual	46
4.2.6 Koefisien Determinasi.....	49
4.2.7 Interpretasi Model Regresi Nonparametrik Spline ...	49

BAB V KESIMPULAN DAN SARAN

5.1 Kesimpulan.....	59
5.2 Saran.....	60

DAFTAR PUSTAKA	61
LAMPIRAN	65
BIODATA PENULIS	97

DAFTAR TABEL

	Halaman
Tabel 2.1 ANOVA untuk Regresi Nonparametrik Spline	12
Tabel 3.1 Variabel Penelitian.....	21
Tabel 3.2 Struktur Data Penelitian.....	23
Tabel 4.1 Statistika Deskriptif Variabel.....	27
Tabel 4.2 Titik Knot dan GCV untuk Satu Knot	37
Tabel 4.3 Titik Knot dan GCV untuk Dua Knot.....	39
Tabel 4.4 Titik Knot dan GCV untuk Tiga Knot	40
Tabel 4.5 Titik Knot dan GCV untuk Kombinasi Knot	42
Tabel 4.6 Perbandingan Nilai GCV Minimum	44
Tabel 4.7 Hasil Pengujian Secara Serentak.....	45
Tabel 4.8 Hasil Pengujian Secara Parsial.....	46
Tabel 4.9 Hasil Pengujian Glejser.....	47

(Halaman ini sengaja dikosongkan)

DAFTAR GAMBAR

	Halaman
Gambar 3.1	Diagram Alir Penelitian 25
Gambar 4.1	Angka Kecelakaan Lalu Lintas di Provinsi Jawa Timur 28
Gambar 4.2	Kepadatan Penduduk di Jawa Timur 29
Gambar 4.3	Pelaku Usia Remaja di Jawa Timur 30
Gambar 4.4	Pelaku Berpendidikan Akhir di Lebih dari SMP di Jawa Timur 31
Gambar 4.5	Rasio Jenis Kelamin Pelaku di Jawa Timur 31
Gambar 4.6	Pelaku Kecelakaan Tanpa SIM di Jawa Timur 32
Gambar 4.7	<i>Scatterplot</i> Antara Angka Laka dengan Kepadatan Penduduk 33
Gambar 4.8	<i>Scatterplot</i> Antara Angka Laka dengan Pelaku Usia Remaja 34
Gambar 4.9	<i>Scatterplot</i> Antara Angka Laka dengan Pelaku Lebih dari SMP 35
Gambar 4.10	<i>Scatterplot</i> Antara Angka Laka dengan Rasio Jenis Kelamin Pelaku 35
Gambar 4.11	<i>Scatterplot</i> Antara Angka Laka dengan Pelaku Tanpa SIM 36
Gambar 4.12	Plot ACF Residual 48
Gambar 4.13	Hasil Uji <i>Kolmogorov Smirnov</i> 48
Gambar 4.14	Persebaran Kab./Kota Berdasarkan Kepadatan Penduduk 51
Gambar 4.15	Persebaran Kab./Kota Berdasarkan Pelaku Usia Remaja 52
Gambar 4.16	Persebaran Kab./Kota Berdasarkan Pelaku Berpendidikan Akhir Lebih dari SMP 55
Gambar 4.17	Persebaran Kab./Kota Berdasarkan Rasio Jenis Kelamin Pelaku 56
Gambar 4.18	Persebaran Kab./Kota Berdasarkan Pelaku Tidak Memiliki SIM 58

(Halaman ini sengaja dikosongkan)

DAFTAR LAMPIRAN

	Halaman
Lampiran 1. Data Angka Kecelakaan Lalu Lintas di Provinsi Jawa Timur dengan Faktor-Faktor yang Mempengaruhi Tahun 2015.....	65
Lampiran 2. Program Pemilihan Titik Knot Optimal dengan Satu Titik Knot Menggunakan Software R.....	67
Lampiran 3. Program Pemilihan Titik Knot Optimal dengan Dua Titik Knot Menggunakan Software R.....	70
Lampiran 4. Program Pemilihan Titik Knot Optimal dengan Tiga Titik Knot Menggunakan <i>Software R</i>	73
Lampiran 5. Program Pemilihan Titik Knot Optimal dengan Kombinasi Titik Knot Menggunakan Software R	76
Lampiran 6. Program Pengujian Parameter Menggunakan Software R	83
Lampiran 7. Uji Glejser Menggunakan Software Minitab...	86
Lampiran 8. Output Nilai GCV dengan Satu Titik Knot.....	87
Lampiran 9. Output Nilai GCV dengan Dua Titik Knot.....	88
Lampiran 10. Output Nilai GCV dengan Tiga Titik Knot	89
Lampiran 11. Output Nilai GCV dengan Kombinasi Titik Knot (1, 3, 3, 3, 3).....	90
Lampiran 12. Output Estimasi Parameter dan Uji Signifikansi.....	91
Lampiran 13. Output Residual	94
Lampiran 14. Surat Keterangan Pengambilan Data	95

(Halaman ini sengaja dikosongkan)

BAB I

PENDAHULUAN

1.1 Latar Belakang

Provinsi Jawa Timur merupakan salah satu provinsi yang memiliki penduduk dengan jumlah yang tinggi, sekaligus menjadi peringkat kedua di Indonesia setelah Jawa Barat pada tahun 2012 hingga 2015. Pada tahun 2012, jumlah penduduk di Jawa Timur mencapai 38.106.590 jiwa, kemudian mengalami peningkatan di tahun 2013 menjadi 38.363.195 jiwa, dan seterusnya hingga tahun 2015 Provinsi Jawa Timur memiliki jumlah penduduk sebanyak 38.847.561 jiwa (BPS, 2015). Hal ini pula yang menjadikan Jawa Timur menjadi salah satu provinsi yang memiliki kepadatan penduduk yang tinggi di Indonesia yakni sebesar 813 jiwa/km² pada tahun 2015 (BPS, 2015). Kepadatan penduduk yang tinggi mengakibatkan Jawa Timur menghadapi berbagai masalah, salah satunya masalah transportasi. Ketergantungan masyarakat Jawa Timur terhadap sektor transportasi, ditambah dengan padatnya penduduk di Jawa Timur menyebabkan banyaknya masalah transportasi yang terjadi di provinsi ini, diantaranya adalah kecelakaan lalu lintas. Data BPS mencatat bahwa selama tiga tahun berturut-turut (2012-2014), Jawa Timur menempati urutan pertama sebagai provinsi dengan jumlah kecelakaan lalu lintas tertinggi di Indonesia. Pada tahun 2012, tercatat jumlah kecelakaan lalu lintas di Jawa Timur sebanyak 24.985 kasus (BPS, 2013). Selanjutnya, pada tahun 2013 menjadi 21.362 kasus dan 19.068 kasus pada tahun 2014 (BPS, 2014).

Menurut WHO dalam Diana (2017), kecelakaan lalu lintas menjadi sepuluh terbesar penyebab kematian di seluruh dunia, dan juga dinilai menjadi pembunuh terbesar ketiga di Indonesia setelah penyakit jantung koroner dan *tuberculosis*. Sebanyak 80% penyebab kecelakaan lalu lintas yang terjadi di jalan raya adalah karena kesalahan manusia (Siahaan dalam Rahmayanti, 2014). Selain itu menurut Rossa dalam Diana (2017), manusia adalah

penyebab utama dalam sebuah peristiwa kecelakaan lalu lintas karena sebagian besar kejadian kecelakaan diawali dengan pelanggaran rambu-rambu lalu lintas. Oleh karena itu, dilakukan penelitian mengenai faktor penyebab terjadinya kecelakaan lalu lintas berdasarkan faktor manusia, mengingat tingginya angka kecelakaan yang terjadi di Jawa Timur. Sehingga dari penelitian ini diharapkan dapat mengetahui hubungan angka kecelakaan lalu lintas di setiap kabupaten/kota di Jawa Timur dengan faktor-faktor yang diduga mempengaruhi agar dapat diupayakan penanggulangannya untuk mengurangi angka kecelakaan lalu lintas yang terjadi.

Salah satu metode yang berfungsi untuk mengetahui model hubungan antar variabel respon dengan variabel prediktor adalah metode regresi. Ada tiga pendekatan dalam metode analisis regresi untuk mengestimasi kurva regresi yaitu pendekatan parametrik, semiparametrik, dan nonparametrik. Pada regresi nonparametrik, data dibiarkan mencari sendiri pola datanya sehingga subjektivitas dari peneliti dapat diminimalisir. Dengan demikian, pendekatan regresi nonparametrik memiliki fleksibilitas yang tinggi (Eubank, 1998). Menurut Wahba (1990) terdapat beberapa pendekatan regresi nonparametrik diantaranya adalah Spline. Pendekatan regresi nonparametrik Spline merupakan model yang mempunyai interpretasi statistik dan visual yang sangat khusus dan sangat baik. Menurut Budiantara (2009) Spline memiliki kemampuan yang sangat baik untuk menangani data yang perilakunya berubah-ubah pada sub interval tertentu. Kelebihan lain dari Spline adalah model cenderung mencari sendiri estimasi data kemanapun pola data tersebut bergerak. Kelebihan ini terjadi karena Spline memiliki titik-titik knot, yaitu titik perpaduan bersama yang menunjukkan terjadinya perubahan pola data.

Penelitian ini menggunakan variabel berdasarkan penelitian-penelitian yang pernah dilakukan sebelumnya, serta didasarkan pada faktor umum penyebab kecelakaan lalu lintas, yaitu

kepadatan penduduk, persentase pelaku usia remaja, persentase pendidikan terakhir pelaku kecelakaan lebih dari SMP, rasio jenis kelamin, dan persentase pelaku tidak memiliki SIM. Berdasarkan data yang digunakan dalam penelitian ini, apabila dilihat dari kurva regresi antara angka kecelakaan lalu lintas di Jawa Timur dengan faktor-faktor yang diduga mempengaruhi tidak membentuk pola tertentu. Kemudian dilihat secara visual, sebagai contoh variabel prediktor yang diduga kuat memengaruhi penyebab kecelakaan lalu lintas salah satunya adalah kepadatan penduduk. Semakin tinggi kepadatan penduduk di suatu daerah, semakin tinggi pula resiko angka kecelakaan lalu lintas yang akan terjadi, namun hal itu belum tentu benar. Tidak semua kabupaten/kota di Jawa Timur akan mengalami hal yang sama, tergantung dari karakteristik yang dimiliki oleh masing-masing daerah, hal ini yang menyebabkan adanya perbedaan pada sub-sub interval (daerah) tertentu. Sehingga pendekatan menggunakan metode regresi nonparametrik Spline cocok digunakan dalam penelitian ini.

Penelitian sebelumnya mengenai kecelakaan lalu lintas telah dilakukan oleh beberapa peneliti diantaranya, Dewi (2016) tentang angka kecelakaan lalu lintas di Jawa Timur dengan metode *Geographically Weighted Regression* menghasilkan kesimpulan faktor-faktor yang mempengaruhi angka kecelakaan lalu lintas di Jawa Timur adalah kepadatan penduduk dan persentase pendidikan terakhir pelaku kecelakaan diatas SMP. Penelitian oleh Rahmayanti (2014) menggunakan metode regresi logistik multinomial menghasilkan kesimpulan bahwa faktor-faktor yang memengaruhi adalah jenis kecelakaan, usia korban, lokasi kecelakaan, jumlah kendaraan yang terlibat, dan fungsi jalan. Penelitian tentang kecelakaan juga pernah dilakukan oleh Gumawang (2015) menggunakan metode pemetaan berjenjang tertimbang menghasilkan kesimpulan variabel yang signifikan terhadap tingkat kecelakaan di kota besar adalah faktor rasio ketersediaan rambu lalu lintas, faktor kecepatan rata-rata, faktor

kepadatan penduduk, trotoar, bahu jalan, pola arus lalu lintas, hambatan samping, dan faktor pelayanan jalan. Penelitian berbasis metode regresi nonparametrik Spline pernah dilakukan oleh Paranoan (2017) mengenai faktor-faktor yang memengaruhi IPM di Sumatera Utara dengan variabel yang berpengaruh adalah angka partisipasi murni SD, SMP, dan SMA, TPAK, dan angka harapan hidup, serta Wulandari (2016) mengenai angka morbiditas di Jawa Timur dengan variabel yang berpengaruh adalah kepadatan penduduk, rata-rata lama sekolah, persentase penduduk miskin, dan Upah Minimum Kabupaten (UMK).

Penelitian mengenai faktor yang mempengaruhi angka kecelakaan lalu lintas di Jawa Timur berdasarkan kabupaten/kota menggunakan metode regresi nonparametrik Spline masih belum ada yang melakukannya hingga saat ini.

1.2 Rumusan Masalah

Provinsi Jawa Timur merupakan salah satu provinsi yang jumlah penduduknya selalu mengalami kenaikan setiap tahunnya. Kepadatan penduduk yang tinggi di Jawa Timur yakni sebesar 813 jiwa/km² memicu permasalahan kecelakaan lalu lintas. Data BPS mencatat bahwa Jawa Timur menjadi provinsi dengan jumlah kecelakaan lalu lintas tertinggi di Indonesia selama tiga tahun berturut-turut. Pada tahun 2012, tercatat jumlah kejadian kecelakaan lalu lintas sebanyak 24.985 kasus (BPS, 2013). Pada tahun 2013 sebanyak 21.362 kasus dan 19.068 kasus pada tahun 2014 di Jawa Timur (BPS, 2014).

Berdasarkan uraian latar belakang di atas, rumusan masalah dalam penelitian ini adalah sebagai berikut.

1. Bagaimana karakteristik angka kecelakaan lalu lintas dan faktor - faktor yang diduga mempengaruhinya di setiap kabupaten/kota di Jawa Timur?
2. Bagaimana pemodelan faktor penyebab angka kecelakaan lalu lintas di Jawa Timur?

1.3 Tujuan Penelitian

Berdasarkan pada latar belakang dan rumusan masalah yang telah diuraikan sebelumnya, tujuan yang ingin dicapai dari penelitian ini adalah sebagai berikut.

1. Mendeskripsikan angka kecelakaan lalu lintas di Jawa Timur dan faktor - faktor yang diduga mempengaruhinya.
2. Memodelkan faktor - faktor yang mempengaruhi angka kecelakaan lalu lintas di Jawa Timur menggunakan regresi nonparametrik Spline *Truncated*.

1.4 Manfaat Penelitian

Dalam penelitian ini diharapkan dapat memberikan manfaat bagi beberapa pihak, diantaranya adalah memberikan informasi untuk pemerintah maupun masyarakat di setiap kabupaten/kota di Jawa Timur dalam mencegah terjadinya kecelakaan lalu lintas, serta menambah pengetahuan mengenai pemodelan pada data angka kecelakaan lalu lintas menggunakan metode regresi nonparametrik Spline *Truncated* sehingga dapat digunakan sebagai bahan acuan dalam penelitian selanjutnya.

1.5 Batasan Masalah

Dalam penelitian ini perlu diberikan batasan permasalahan agar penelitian yang dikerjakan lebih fokus dan sesuai dengan rentang waktu yang direncanakan. Batasan masalah pada penelitian ini adalah sebagai berikut.

1. Data kecelakaan lalu lintas beserta faktor – faktor yang mempengaruhinya menggunakan data tahun 2015.
2. Usia remaja yang digunakan 10-15 tahun.
3. Faktor penyebab angka kecelakaan lalu lintas hanya meneliti pada faktor manusia.
4. Metode yang digunakan dalam pemodelan adalah Spline *Truncated* linear dengan satu knot, dua knot, tiga knot, dan kombinasi knot.

5. Pemilihan titik knot optimal menggunakan metode *Generalized Cross Validation* (GCV).
6. Panjang jalan dalam penelitian ini menggunakan jalan kabupaten/kota.

BAB II

TINJAUAN PUSTAKA

2.1 Statistika Deskriptif

Statistika deskriptif merupakan suatu metode yang berkaitan dengan pengumpulan dan penyajian data sehingga memberikan informasi yang berguna (Walpole, 1993). Statistika deskriptif yang digunakan memuat beberapa hal sebagai berikut:

a. Rata-rata

Rata-rata merupakan suatu ukuran pusat data bila data tersebut diurutkan dari yang terkecil hingga yang terbesar ataupun sebaliknya. Secara umum, rata-rata diberikan oleh rumus pada persamaan 2.1, dimana \bar{x} adalah rata-rata dan n menunjukkan banyaknya data.

$$\bar{x} = \frac{\sum_{i=1}^n x_i}{n} \quad (2.1)$$

b. Standar Deviasi

Standar deviasi (s) digunakan untuk mengetahui seberapa jauh suatu pengamatan menyebar dari rata-ratanya yang diberikan oleh rumus sebagai berikut.

$$s = \sqrt{\frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n - 1}} \quad (2.2)$$

c. Diagram Batang

Suatu bentuk penyajian data berupa grafik yang berbentuk persegi panjang dengan lebar yang sama dan digunakan pada data numerik.

2.2 Analisis Regresi

Analisis regresi merupakan sebuah metode statistika yang memberikan penjelasan tentang pola hubungan (model) antara

dua variabel atau lebih (Drapper dan Smith, 1992). Model regresi linear berganda diberikan oleh rumus sebagai berikut.

$$y_i = \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + \dots + \beta_p x_{pi} + \varepsilon_i, i = 1, 2, \dots, n \quad (2.3)$$

Berdasarkan model regresi di atas, y_i merupakan variabel respon dan x merupakan variabel prediktor. Analisis regresi terdiri dari tiga pendekatan dalam mengestimasi model yaitu regresi parametrik, semiparametrik, dan regresi nonparametrik. Regresi parametrik digunakan jika kurva regresi mengikuti pola tertentu atau membentuk pola data yang jelas seperti linier, kuadratik, dan kubik. Regresi nonparametrik digunakan apabila kurva regresi tidak diketahui polanya, sedangkan regresi semiparametrik digunakan apabila terdapat komponen parametrik dan nonparametrik (Budiantara, 2009).

2.3 Regresi Nonparametrik Spline

Regresi nonparametrik merupakan sebuah metode regresi yang tidak terikat asumsi bentuk kurva regresi tertentu (Budiantara, 2001). Pendekatan regresi nonparametrik digunakan untuk menyelesaikan pola data antara variabel respon dengan variabel prediktor yang tidak membentuk pola tertentu. Fungsi regresi diasumsikan *smooth* (mulus) yang artinya termuat dalam suatu ruang fungsi tertentu. Berdasarkan pendekatan nonparametrik, data diharapkan mencari sendiri bentuk estimasinya tanpa dipengaruhi oleh faktor subjektivitas dari peneliti, dengan demikian pendekatan regresi nonparametrik memiliki fleksibilitas yang tinggi (Eubank, 1988). Bentuk model regresi nonparametrik secara umum adalah sebagai berikut.

$$y_i = f(x_i) + \varepsilon_i \quad (2.4)$$

dimana,

y_i : variabel respon ke- i , $i = 1, 2, \dots, n$

f : kurva regresi yang diasumsikan bentuknya tidak diketahui

x_i : variabel prediktor ke- i

ε_i : *error* random ke- i yang diasumsikan identik, independen, dan berdistribusi normal.

Menurut Wahba (1990) terdapat beberapa pendekatan regresi nonparametrik diantaranya adalah Spline. Spline memiliki kemampuan yang sangat baik untuk menangani data yang perilakunya berubah-ubah pada sub interval tertentu, serta memiliki kelebihan untuk cenderung mencari sendiri estimasi data kemana pola data tersebut bergerak (Budiantara, 2009). Kelebihan ini terjadi karena Spline memiliki titik-titik knot K_1, K_2, \dots, K_r . Spline merupakan potongan-potongan polinomial yang memiliki sifat tersegmen pada titik knot. Salah satu kelemahan fungsi polinomial adalah bersifat global, sehingga dikembangkan suatu fungsi yang dapat mengatasi kelemahan polinomial yaitu fungsi Spline *Truncated* yang mempertahankan fungsi polinomial. Fungsi Spline *Truncated* univariabel dengan derajat m diberikan oleh persamaan sebagai berikut.

$$f(x_i) = \sum_{j=0}^m \beta_j x_i^j + \sum_{k=1}^r \beta_{m+k} (x_i - K_k)_+^m \quad (2.5)$$

Sehingga model regresi nonparametrik Spline univariabel berdasarkan persamaan 2.4 dapat dituliskan sebagai berikut.

$$y_i = \sum_{j=0}^m \beta_j x_i^j + \sum_{k=1}^r \beta_{m+k} (x_i - K_k)_+^m + \varepsilon_i, \quad i = 1, 2, \dots, n \quad (2.7)$$

Fungsi $(x - K_k)_+^m$ adalah fungsi potongan (*Truncated*) yang diberikan oleh,

$$(x_i - K_k)_+^m = \begin{cases} (x_i - K_k)^m, & x_i \geq K_k \\ 0 & , x_i < K_k \end{cases} \quad (2.8)$$

dimana,

β_j : parameter model polinomial, $j = 1, 2, \dots, m$

x_i : variabel prediktor, $i = 1, 2, \dots, n$

β_{m+k} : parameter pada komponen *Truncated*, $k = 1, 2, \dots, r$

r : banyaknya knot

K_k : titik knot, titik yang menunjukkan perubahan pola data

Apabila terdapat lebih dari satu prediktor, maka regresi Spline yang digunakan adalah regresi Spline multivariabel. Misalkan y_i merupakan variabel respon, $i = 1, 2, \dots, n$, sedangkan $x_{1i}, x_{2i}, \dots, x_{pi}$ adalah variabel prediktor yang pola hubungannya dengan variabel respon tidak diketahui polanya. Hubungan keduanya dapat dimodelkan dengan regresi nonparametrik multivariabel sebagai berikut.

$$\begin{aligned} y_i &= f(x_{1i}) + f(x_{2i}) + \dots + f(x_{pi}) + \varepsilon_i \\ &= \sum_{u=1}^p f(x_{ui}) + \varepsilon_i, \quad i = 1, 2, \dots, n \end{aligned} \quad (2.9)$$

dimana, n menunjukkan banyak observasi dan p menunjukkan banyaknya variabel prediktor. Pada regresi nonparametrik Spline, fungsi $f(x_{ui})$ dapat ditulis sebagai berikut.

$$f(x_{ui}) = \sum_{j=0}^m \beta_{uj} x_{ui}^j + \sum_{k=1}^r \beta_{u(m+k)} (x_{ui} - K_{uk})_+^m \quad (2.10)$$

Sehingga model regresi nonparametrik Spline multivariabel dapat dituliskan sebagai berikut.

$$\begin{aligned} y_i &= \sum_{u=1}^p \left(\sum_{j=0}^m \beta_{uj} x_{ui}^j + \sum_{k=1}^r \beta_{u(m+k)} (x_{ui} - K_{uk})_+^m \right) + \varepsilon \\ &= \sum_{u=1}^p \sum_{j=0}^m \beta_{uj} x_{ui}^j + \sum_{u=1}^p \sum_{k=1}^r \beta_{u(m+k)} (x_{ui} - K_{uk})_+^m + \varepsilon \end{aligned} \quad (2.11)$$

Fungsi $(x_{ui} - K_{uk})_+^m$ adalah fungsi potongan (*Truncated*) yang diberikan oleh,

$$(x_{ui} - K_{uk})_+^m = \begin{cases} (x_{ui} - K_{uk})^m, & x_{ui} \geq K_{uk} \\ 0, & x_{ui} < K_{uk} \end{cases} \quad (2.12)$$

2.4 Pemilihan Titik Knot Optimal

Titik knot merupakan titik perpaduan bersama yang menunjukkan perubahan pola perilaku kurva fungsi Spline pada selang yang berbeda (Hardle, 1990). Terdapat beberapa metode yang digunakan dalam pemilihan titik knot yang optimal dengan pendekatan Spline, salah satunya adalah metode *Generalized Cross Validation* (GCV). Menurut Budiantara (2006), titik-titik knot optimal diperoleh dari nilai GCV paling minimum, secara umum dapat dituliskan sebagai berikut.

$$GCV(K_1, K_2, \dots, K_r) = \frac{MSE(K_1, K_2, \dots, K_r)}{(n^{-1} \text{trace}[\mathbf{I} - \mathbf{A}(K_1, K_2, \dots, K_r)])^2} \quad (2.13)$$

dimana,

K_1, K_2, \dots, K_r : titik knot pertama hingga titik knot ke- r

\mathbf{I} : matriks identitas

n : jumlah pengamatan

$$\mathbf{A}(K_1, K_2, \dots, K_r) = \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'$$

$$MSE(K_1, K_2, \dots, K_r) = n^{-1} \sum_{i=1}^n (y_i - \hat{y}_i)^2 \quad (2.14)$$

2.5 Pengujian Parameter Model

Pengujian parameter model dilakukan untuk mengetahui apakah variabel memberikan pengaruh signifikan terhadap model. Pengujian ini terdiri dari pengujian secara serentak dan pengujian secara parsial.

2.5.1 Pengujian Secara Serentak

Uji serentak merupakan uji signifikansi model secara keseluruhan atau untuk mengetahui apakah semua variabel prediktor dalam model secara bersama-sama memberikan pengaruh terhadap variabel respon. Berikut ini adalah hipotesis yang digunakan.

$$H_0 : \beta_1 = \beta_2 = \dots = \beta_{m+r} = 0$$

$$H_1 : \text{minimal ada satu } \beta_g \neq 0, \quad g = 1, 2, \dots, m+r$$

dengan $m + r$ merupakan banyaknya parameter dalam model regresi nonparametrik Spline kecuali β_0 . Statistik uji menggunakan uji F sebagai berikut.

$$F_{hitung} = \frac{MSR}{MSE} \quad (2.15)$$

dimana, MSR (*Mean Square Regresi*) adalah hasil bagi antara jumlah kuadrat regresi dengan df regresi, sedangkan MSE (*Mean Square Error*) adalah hasil bagi dari jumlah kuadrat error dengan df error yang disajikan pada Tabel 2.1 menggunakan analisis ragam (ANOVA).

Tabel 2.1 ANOVA untuk Regresi Nonparametrik Spline

Sumber Variasi	Derajat Bebas (df)	Jumlah Kuadrat (Sum Square)	Kuadrat Tengah (Mean Square)	F
Regresi	$m+r$	$\sum_{i=1}^n (\hat{y}_i - \bar{y})^2$	$\frac{\sum_{i=1}^n (\hat{y}_i - \bar{y})^2}{m+r}$	$\frac{MSR}{MSE}$
Error	$n-(m+r)-1$	$\sum_{i=1}^n (y_i - \hat{y}_i)^2$	$\frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{n-(m+r)-1}$	
Total	$n-1$	$\sum_{i=1}^n (y_i - \bar{y})^2$	-	-

Daerah penolakan yang digunakan adalah H_0 ditolak jika $F_{hitung} > F_{\alpha, (m+r), (n-(m+r)-1)}$. Apabila keputusan menolak H_0 maka dapat disimpulkan minimal terdapat satu parameter pada model regresi Spline yang signifikan terhadap model.

2.5.2 Pengujian Secara Parsial

Apabila dalam pengujian secara serentak didapatkan kesimpulan minimal terdapat satu parameter pada model regresi Spline yang signifikan, maka perlu dilanjutkan uji secara parsial. Variabel prediktor dikatakan berpengaruh terhadap variabel

respon apabila terdapat minimal satu parameter yang signifikan. Hipotesis yang digunakan adalah sebagai berikut:

$$H_0 : \beta_g = 0$$

$$H_1 : \beta_g \neq 0, \quad g = 1, 2, \dots, m+r$$

Statistik uji yang digunakan adalah uji t sebagai berikut.

$$t_{hitung} = \frac{\hat{\beta}_g}{se(\hat{\beta}_g)} \quad (2.16)$$

dimana,

$\hat{\beta}_g$: penaksir parameter β_g

$se(\hat{\beta}_g)$: standart error dari $\hat{\beta}_g$

Nilai $se(\hat{\beta}_g)$ didapatkan dari $\sqrt{\text{var}(\hat{\beta}_g)}$, $\text{var}(\hat{\beta}_g)$ merupakan elemen diagonal ke- g dari matriks sebagai berikut :

$$\begin{aligned} \text{var}(\hat{\beta}) &= \text{var}[(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y}] \\ &= (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}' \text{var}(\mathbf{y})[(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}]' \\ &= \sigma^2(\mathbf{X}'\mathbf{X})^{-1} \end{aligned}$$

Nilai σ^2 didekati dengan nilai MSE. Daerah penolakan yang digunakan adalah H_0 ditolak jika nilai $|t_{hitung}| > t_{\alpha/2; (n-(m+r)-1)}$. Kesimpulan yang diperoleh jika tolak H_0 adalah parameter berpengaruh signifikan terhadap model (Drapper & Smith, 1992).

2.6 Pengujian Asumsi Residual

Apabila model terbaik dari regresi Spline telah didapatkan, maka perlu dilakukan pengujian asumsi residual untuk mengetahui apakah residual yang dihasilkan telah memenuhi asumsi yaitu identik, independen dan berdistribusi normal.

2.6.1 Uji Asumsi Identik

Pengujian ini dilakukan untuk mengetahui homogenitas varians residual model regresi. Apabila varians antar residual

tidak homogen maka terjadi kasus heteroskedastisitas. Heteroskedastisitas dapat dideteksi dengan uji *Glejser*. Uji ini dilakukan dengan meregresikan nilai mutlak residual dengan seluruh variabel prediktor. Hipotesis yang digunakan dalam pengujian ini adalah sebagai berikut.

$$H_0 : \sigma_1^2 = \sigma_2^2 = \dots = \sigma_n^2 = \sigma^2 \text{ (residual identik)}$$

$$H_1 : \text{minimal terdapat satu } \sigma_i^2 \neq \sigma^2, i=1,2,\dots,n \text{ (residual tidak identik)}$$

Statistik uji yang digunakan adalah sebagai berikut.

$$F_{hitung} = \frac{\sum_{i=1}^n (|\hat{e}_i| - |\bar{e}|)^2 / (v-1)}{\sum_{i=1}^n (|e_i| - |\hat{e}_i|)^2 / (n-v)} \quad (2.17)$$

Daerah penolakan yang digunakan adalah tolak H_0 apabila $F_{hitung} > F_{\alpha, (v-1), (n-v)}$, dimana nilai v merupakan banyaknya parameter model *Glejser*. Jika H_0 ditolak maka dapat disimpulkan bahwa terdapat kasus heteroskedastisitas sehingga asumsi identik tidak terpenuhi (Gujarati, 2004).

2.6.2 Asumsi Independen

Pengujian asumsi independen dilakukan untuk mengetahui apakah terdapat autokorelasi atau tidak pada residual. Ada beberapa cara untuk mendeteksi autokorelasi diantaranya adalah melihat plot ACF (*Autocorrelation Function*) dari residual. Untuk mendapatkan nilai ACF digunakan rumus sebagai berikut.

$$\hat{\rho}_k = \frac{\sum_{t=k+1}^{n-k} (e_t - \bar{e})(e_{t-k} - \bar{e})}{\sum_{t=1}^n (e_t - \bar{e})^2}, k=1,2,3,\dots \quad (2.18)$$

dimana, $\hat{\rho}_k$ merupakan korelasi antara e_t dan e_{t-k} , dan k adalah lag ke- k . Interval konfidensi $(1-\alpha)100\%$ untuk autokorelasi ρ_k diberikan oleh persamaan 2.19.

$$-t_{\alpha/2, n-1} se(\hat{\rho}_k) < \rho_k < t_{1-\alpha/2, n-1} se(\hat{\rho}_k) \quad (2.19)$$

dengan,

$$se(\hat{\rho}_k) = \sqrt{\frac{1 + 2 \sum_{k=1}^{k-1} \hat{\rho}_k^2}{n}} \quad (2.20)$$

Apabila terdapat ACF (ρ_k) yang keluar dari interval konfidensi, maka diindikasikan adanya kasus autokorelasi antar residual. Sebaliknya, jika semua nilai ACF berada di dalam batas interval maka tidak terdapat kasus autokorelasi atau asumsi independen terpenuhi (Wei, 2006).

2.6.3 Uji Asumsi Berdistribusi Normal

Uji asumsi distribusi normal dilakukan untuk mengetahui apakah residual berdistribusi normal atau tidak. Secara visual pengujian asumsi distribusi normal dapat dilakukan dengan *normal probability plot*. Cara lain dapat dilakukan pengujian *Kolmogorov Smirnov* dengan hipotesis sebagai berikut.

$H_0 : F(x) = F_0(x)$ (residual berdistribusi normal)

$H_1 : F(x) \neq F_0(x)$ (residual tidak berdistribusi normal)

Statistik uji yang digunakan dalam uji *Kolmogorov Smirnov* adalah sebagai berikut.

$$D = \sup_x |S(x) - F_0(x)| \quad (2.21)$$

dimana $S(x)$ adalah fungsi peluang kumulatif yang dihitung dari data sampel, sedangkan $F_0(x)$ adalah fungsi peluang kumulatif distribusi normal, serta $F(x)$ adalah fungsi distribusi yang belum diketahui. Daerah penolakan pada uji *Kolmogorov Smirnov* yaitu tolak H_0 apabila $|D| > q_{(1-\alpha, n)}$ atau $p\text{-value} < \alpha$, dimana nilai $q_{(1-\alpha, n)}$ didapatkan dari tabel *Kolmogorov Smirnov*. Keputusan yang didapatkan jika tolak H_0 adalah residual tidak berdistribusi normal (Daniel, 1989).

2.7 Kriteria Pemilihan Model Terbaik

Salah satu kriteria yang digunakan untuk menentukan model regresi terbaik yaitu menggunakan koefisien determinasi (R^2). Koefisien determinasi adalah proporsi keragaman atau variansi total nilai variabel y yang dapat dijelaskan oleh variabel x (Drapper & Smith, 1992). Semakin tinggi nilai R^2 yang dihasilkan dari suatu model, maka semakin baik pula variabel prediktor dalam model regresi dalam menjelaskan variabilitas variabel respon. Berikut merupakan rumus dari koefisien determinasi (R^2).

$$R^2 = \frac{SSR}{SST} = \frac{\sum_{i=1}^n (\hat{y}_i - \bar{y})^2}{\sum_{i=1}^n (y_i - \bar{y})^2} \quad (2.22)$$

2.8 Angka Kecelakaan Lalu Lintas

Berdasarkan Undang-Undang RI No.22 Tahun 2009 tentang lalu lintas dan angkutan jalan, kecelakaan lalu lintas merupakan suatu peristiwa di jalan yang tidak diduga dan tidak disengaja melibatkan kendaraan dengan atau tanpa pengguna jalan lain yang mengakibatkan korban manusia maupun kerugian harta benda. Kasus kecelakaan sulit diminimalisasi dan cenderung meningkat seiring pertambahan panjang jalan (Hobbs, 1995).

Menurut Pignataro dalam Rahmadi (2011), angka kecelakaan (*accident rate*) biasanya digunakan untuk mengukur tingkat kecelakaan pada satu satuan panjang jalan. Dalam penelitian ini angka kecelakaan lalu lintas digunakan sebagai alat penentu untuk menyatakan tingkat kecelakaan lalu lintas pada satu satuan segmen jalan di tiap kabupaten/kota di Jawa Timur. Panjang jalan yang digunakan dalam penelitian ini adalah jalan kabupaten/kota se-Jawa Timur. Jalan kabupaten/kota adalah jalan lokal dalam sistem jaringan jalan primer yang tidak termasuk pada jalan nasional dan jalan provinsi, yang menghubungkan ibu kota kabupaten dengan ibu kota kecamatan, antar ibu kota kecamatan,

ibu kota kabupaten dengan pusat kegiatan lokal, antar pusat kegiatan lokal, serta jalan umum dalam sistem jaringan jalan sekunder dalam wilayah kabupaten, dan jalan strategis kabupaten (BPS, 2013). Penelitian mengenai angka kecelakaan lalu lintas menggunakan panjang jalan kabupaten/kota di Provinsi Jawa Timur pernah dilakukan oleh Dewi (2016).

Berdasarkan data BPS, diketahui bahwa jalan kabupaten/kota merupakan bagian terbesar yaitu 415.788 kilometer atau 81,85% dari total panjang jalan di Indonesia, namun yang memiliki kondisi baik hanya 39,97%, kondisi sedang 22,08%, rusak 22,33% dan sisanya dalam kondisi rusak berat.

2.9 Faktor Penyebab Kecelakaan Lalu Lintas

Secara umum, kecelakaan lalu lintas disebabkan oleh berbagai faktor diantaranya adalah faktor manusia, faktor lingkungan, faktor jalan, dan faktor kendaraan (Warpani dalam Rahmayanti, 2014). Sebanyak 80% penyebab kecelakaan lalu lintas yang terjadi di jalan raya adalah karena kesalahan manusia (Siahaan dalam Rahmayanti, 2014). Selain itu menurut Rossa dalam Rahmayanti (2014), manusia merupakan penyebab utama dalam sebuah peristiwa kecelakaan lalu lintas karena sebagian besar kejadian kecelakaan diawali dengan pelanggaran rambu-rambu lalu lintas. Sehingga dalam penelitian ini hanya ingin memfokuskan faktor penyebab kecelakaan yang disebabkan oleh manusia. Berdasarkan penelitian-penelitian sebelumnya, faktor-faktor yang berkaitan dengan manusia sebagai penyebab kecelakaan lalu lintas adalah sebagai berikut.

a. Kepadatan Penduduk

Kepadatan penduduk dibagi menjadi tiga jenis, yaitu Kepadatan Penduduk Kasar, Kepadatan Penduduk Fisiologis, dan Kepadatan Penduduk Agraris. Kepadatan penduduk kasar merupakan ukuran persebaran penduduk umum digunakan, karena menunjukkan banyaknya jumlah penduduk untuk setiap kilometer persegi luas wilayah. Pertumbuhan penduduk

yang semakin bertambah setiap tahunnya dan tidak diikuti dengan peningkatan luas wilayah menyebabkan terjadinya kepadatan penduduk yang tinggi. Peningkatan jumlah penduduk dengan sikap individu yang kurang peduli terhadap lingkungan yang aman dan tertib menjadi salah satu penyebab terjadinya kecelakaan lalu lintas, sehingga banyak pelanggaran lalu lintas yang dilakukan oleh pengemudi (Triana, dkk., 2013).

Gumawang (2015) pernah meneliti tentang kecelakaan lalu lintas di kota besar dan didapatkan hasil bahwa variabel yang signifikan terhadap angka kecelakaan di kota besar adalah faktor rasio ketersediaan rambu lalu lintas, faktor kecepatan rata-rata, faktor kepadatan penduduk, trotoar, bahu jalan, pola arus lalu lintas, hambatan samping, dan faktor pelayanan jalan. Berdasarkan penelitian tersebut terlihat bahwa kepadatan penduduk menjadi salah satu penyebab yang memberikan kontribusi terhadap angka kecelakaan lalu lintas.

b. Persentase Pelaku Kecelakaan Usia Remaja

Menurut Maulina dalam Dewi (2016), usia merupakan salah satu faktor internal yang berpengaruh terhadap sikap agresif dan emosional seorang pengemudi. Menurut BPS (2015), usia masa remaja adalah peralihan dari masa anak ke masa dewasa yang mengalami perkembangan semua aspek/fungsi. Beberapa ahli menggunakan batasan usia yang berbeda untuk mengelompokkan usia remaja. Menurut Monks, dkk dalam BPS (2015), masa remaja dibedakan menjadi empat golongan yaitu,

- (i) masa pra-remaja: $10 \leq \text{usia} \leq 12$ tahun
- (ii) masa remaja awal: $12 < \text{usia} \leq 15$ tahun
- (iii) masa remaja pertengahan: $15 < \text{usia} \leq 18$ tahun
- (iv) masa remaja akhir: $18 < \text{usia} \leq 21$ tahun.

Dalam penelitian ini pelaku usia remaja yang digunakan adalah pelaku pada rentang usia 10-15 tahun yaitu pada masa pra remaja hingga remaja awal, dimana pada kelompok umur tersebut merupakan pengemudi pemula dengan tingkat emosi yang belum

stabil dan belum berhati-hati dalam berkendara. Selain itu, belum adanya kepemilikan SIM (Surat Izin Mengemudi) pada kelompok umur tersebut sehingga tidak mengetahui dengan jelas edukasi berkendara dengan aman yang sesuai dengan hukum berlaku.

Menurut Sabey dalam Kartika (2009), orang yang berusia muda lebih sering terlibat dalam suatu kecelakaan lalu lintas, baik sebagai pejalan kaki maupun pengemudi dibandingkan dengan orang yang berusia lebih tua. Penelitian yang pernah dilakukan oleh Goodwin,dkk. (2014) menunjukkan bahwa usia remaja menyumbangkan angka kecelakaan tertinggi di Texas. Data dari Korlantas Polri pada tahun 2015 menyatakan bahwa pelajar masuk dalam urutan kedua terbanyak berdasarkan klasifikasi pelaku kecelakaan, data ini semakin mempertegas bahwa usia muda/remaja memang rawan menjadi penyebab kecelakaan (Marta, 2017).

c. Persentase Pendidikan Terakhir Pelaku Lebih dari SMP

Karakteristik pengemudi kendaraan juga dipengaruhi oleh pendidikan yang ditamatkan oleh pengemudi tersebut. Menurut Wahyuningtyas dalam Dewi (2016) menyatakan bahwa terdapat perbedaan yang signifikan antara pelaku kecelakaan dengan pendidikan yang ditamatkan SD, SMP, SMA atau Perguruan Tinggi yang dipengaruhi oleh sikap disiplin berlalu lintas yang berbeda. Selain itu tingkat pendidikan sangat berpengaruh terhadap program peningkatan pengetahuan secara langsung dan secara tidak langsung terhadap perilaku berkendara. Pada umumnya seseorang yang berpendidikan rendah mempunyai ciri sulit untuk diajak bekerja sama dan kurang terbuka terhadap pembaharuan, hal ini disebabkan masih adanya nilai-nilai lama yang mereka anut selama ini, serta pengemudi yang telah menyelesaikan pendidikan SMP dianggap memiliki etika berkendara yang lebih baik (Hubdat dalam Dewi, 2016).

Penelitian yang pernah dilakukan oleh Dewi (2016) mengenai faktor-faktor penyebab angka kecelakaan lalu lintas di setiap kabupaten/kota di Jawa Timur menggunakan metode GWR

(*Geographically Weighted Regression*) dan didapatkan hasil bahwa variabel yang mempengaruhi angka kecelakaan lalu lintas adalah kepadatan penduduk dan persentase pendidikan terakhir pelaku kecelakaan lebih dari SMP. Berdasarkan penelitian tersebut terlihat bahwa persentase pendidikan terakhir pelaku kecelakaan di atas SMP menjadi salah satu faktor yang memberikan kontribusi terhadap angka kecelakaan lalu lintas di Jawa Timur.

d. Rasio Jenis Kelamin Pelaku

Karakteristik pengemudi terhadap angka kecelakaan lalu lintas juga ditentukan oleh jenis kelamin pengendara. Menurut Hubdat dalam Dewi (2016), angka kematian akibat kecelakaan lalu lintas pada laki-laki lebih tinggi daripada perempuan. Hal ini dikarenakan jenis kelamin perempuan sebagai pengguna sepeda motor jumlahnya lebih sedikit dibandingkan dengan jumlah pengguna sepeda motor berjenis kelamin laki-laki.

e. Persentase pelaku kecelakaan tidak memiliki SIM

Menurut Notosiswoyo dalam Anwar (2017), pengetahuan pada pengendara merupakan hubungan nyata dalam pencegahan kecelakaan. Pengetahuan yang dimaksud yaitu mengenai larangan berkendara bagi pengendara yang tidak memiliki SIM (Surat Izin Mengemudi), serta melanggar tata tertib lalu lintas dalam berkendara yang benar. Salah satu penyebab dari kecelakaan lalu lintas yang melibatkan generasi muda ialah karena banyak yang mengemudi tapi tidak memiliki SIM. Tidak memiliki SIM dianggap berpotensi menyebabkan pelanggaran lalu lintas yang dapat menyebabkan kecelakaan saat berkendara. Begitu pula menurut Bakharuddin dalam Satria (2014), salah satu penyebab tingginya angka kecelakaan dikarenakan tingginya jumlah pengemudi yang belum memiliki SIM.

BAB III

METODOLOGI PENELITIAN

3.1 Sumber Data

Data yang digunakan dalam penelitian ini adalah data sekunder yang diambil dari dua sumber yaitu Kepolisian Daerah (Polda) Jawa Timur dan Badan Pusat Statistik (BPS) Provinsi Jawa Timur. Data yang digunakan adalah tahun 2015 dengan unit penelitian sebanyak 38 kabupaten/kota di Provinsi Jawa Timur.

3.2 Variabel Penelitian

Variabel yang akan diteliti terdiri dari variabel respon dan variabel prediktor dimana skala data seluruh variabel adalah rasio. Variabel respon yang digunakan adalah angka kecelakaan lalu lintas di Jawa Timur. Sedangkan variabel prediktor yang diduga mempengaruhi angka kecelakaan lalu lintas di Jawa Timur dalam penelitian ini diperoleh dari penelitian-penelitian sebelumnya yang disajikan pada Tabel 3.1.

Tabel 3.1 Variabel Penelitian

Variabel	Keterangan	Sumber
y	Angka kecelakaan lalu lintas	Polda Jatim dan BPS
x_1	Kepadatan penduduk	Badan Pusat Statistik
x_2	Persentase pelaku usia remaja	Polda Jatim
x_3	Persentase pendidikan terakhir pelaku lebih dari SMP	Polda Jatim
x_4	Rasio jenis kelamin pelaku	Badan Pusat Statistik
x_5	Persentase pelaku kecelakaan tidak memiliki SIM	Polda Jatim

3.3 Definisi Operasional

Penjelasan singkat mengenai variabel yang digunakan dalam penelitian ini adalah sebagai berikut.

1. y : Angka kecelakaan lalu lintas

Angka kecelakaan lalu lintas digunakan sebagai alat penentu untuk menyatakan tingkat kecelakaan lalu lintas pada satu satuan segmen jalan.

$$AR = \frac{A}{L} \times 100$$

AR : Angka kecelakaan total per 100 kilometer

A : Total kecelakaan yang terjadi pada tahun tertentu

L : Panjang dari bagian jalan yang di kontrol (km)

2. x_1 : Kepadatan penduduk

Banyaknya jumlah penduduk per satuan luas (BPS). Secara matematis dituliskan sebagai berikut.

$$x_1 = \frac{\text{Jumlah penduduk suatu wilayah (jiwa)}}{\text{Luas wilayah (km}^2\text{)}}$$

3. x_2 : Persentase pelaku usia remaja

Usia remaja yang digunakan adalah rentang 10-15 tahun. Secara matematis, perhitungan persentase pelaku usia remaja diperoleh dengan cara sebagai berikut.

$$x_2 = \frac{\text{Jumlah pelaku kecelakaan berusia 10-15 tahun}}{\text{Jumlah pelaku kecelakaan seluruh usia}} \times 100\%$$

4. x_3 : Persentase pendidikan terakhir pelaku kecelakaan lebih dari SMP

Pendidikan terakhir pelaku kecelakaan lebih dari SMP merupakan pelaku kecelakaan yang berpendidikan terakhir SMA dan perguruan tinggi. Secara matematis, perhitungan variabel ini diperoleh dengan cara sebagai berikut.

$$x_3 = \frac{\text{Jumlah pelaku kecelakaan berpendidikan terakhir lebih dari SMP}}{\text{Jumlah pelaku kecelakaan seluruh pendidikan terakhir}} \times 100\%$$

5. x_4 : Rasio jenis kelamin pelaku

Secara matematis, perhitungan variabel ini dapat dirumuskan dengan cara sebagai berikut.

$$x_4 = \frac{PK_W}{PK_L} \times 100$$

PK_W : Jumlah pelaku kecelakaan berjenis kelamin wanita

PK_L : Jumlah pelaku kecelakaan berjenis kelamin laki-laki

6. x_5 : Persentase pelaku kecelakaan tidak memiliki SIM

Secara matematis, perhitungan variabel ini dapat dirumuskan dengan cara sebagai berikut.

$$x_5 = \frac{\text{Jumlah pelaku kecelakaan yang tidak memiliki SIM}}{\text{Jumlah pelaku kecelakaan seluruh kepemilikan SIM}} \times 100\%$$

3.4 Struktur Data

Berikut ini adalah struktur data yang digunakan dalam penelitian ini dari variabel respon dan variabel prediktor diuraikan dalam Tabel 3.2.

Tabel 3.2 Struktur Data Penelitian

Kab/kota	y	x_1	x_2	x_3	x_4	x_5
1	y_1	$x_{1,1}$	$x_{2,1}$	$x_{3,1}$	$x_{4,1}$	$x_{5,1}$
2	y_2	$x_{1,2}$	$x_{2,2}$	$x_{3,2}$	$x_{4,2}$	$x_{5,2}$
3	y_3	$x_{1,3}$	$x_{2,3}$	$x_{3,3}$	$x_{4,3}$	$x_{5,3}$
\vdots	\vdots	\vdots	\vdots	\vdots	\vdots	\vdots
38	y_{38}	$x_{1,38}$	$x_{2,38}$	$x_{3,38}$	$x_{4,38}$	$x_{5,38}$

3.5 Langkah Analisis

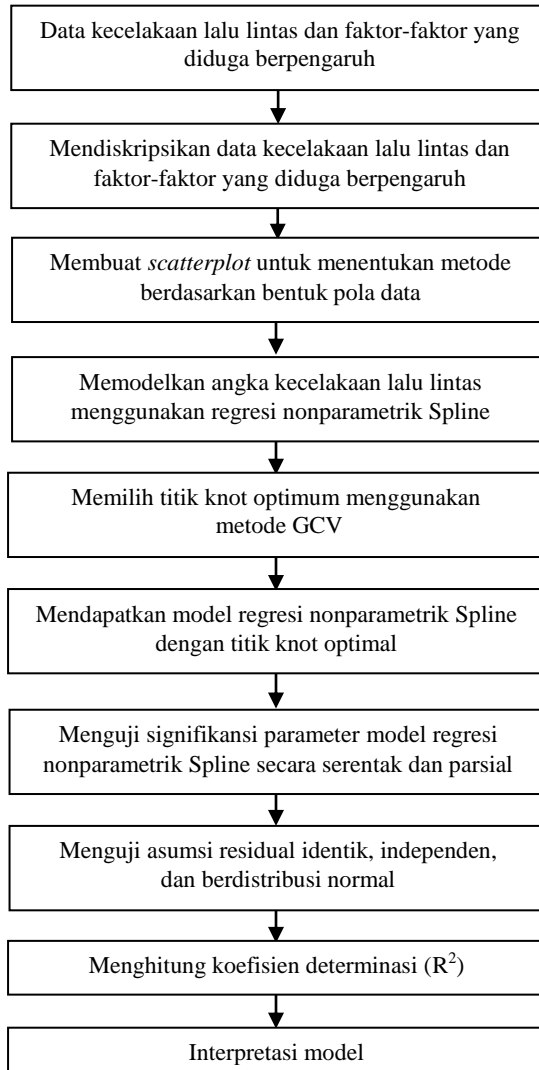
Mengacu pada tujuan penelitian, langkah analisis yang dilakukan dalam penelitian ini adalah sebagai berikut.

1. Mendeskripsikan karakteristik angka kecelakaan lalu lintas di Provinsi Jawa Timur beserta faktor-faktor yang diduga mempengaruhinya.

- a. Melakukan analisis statistika deskriptif menggunakan rata-rata, standar deviasi, dan diagram batang pada seluruh variabel.
 - b. Membuat *scatterplot* antara variabel respon yaitu angka kecelakaan lalu lintas di Jawa Timur dengan masing-masing variabel prediktor untuk mengetahui bentuk pola data. Apabila bentuk pola data tidak mengikuti pola tertentu maka metode yang digunakan adalah regresi nonparametrik Spline.
2. Memodelkan faktor-faktor yang diduga mempengaruhi angka kecelakaan lalu lintas di Provinsi Jawa Timur menggunakan regresi nonparametrik Spline.
 - a. Memodelkan angka kecelakaan lalu lintas menggunakan regresi nonparametrik Spline dengan satu titik knot, dua titik knot, tiga titik knot, dan kombinasi titik knot.
 - b. Memilih titik knot optimal berdasarkan nilai *Generalized Cross Validation* (GCV) yang paling minimum.
 - c. Mendapatkan model regresi nonparametrik Spline terbaik dengan titik knot optimal.
 - d. Melakukan pengujian signifikansi parameter model regresi nonparametrik Spline secara serentak dan parsial sesuai dengan persamaan 2.15 dan persamaan 2.16.
 - e. Melakukan uji asumsi residual identik, independen, dan berdistribusi normal (IIDN) berdasarkan model regresi nonparametrik Spline.
 - f. Menguji kebaikan model dengan menghitung koefisien determinasi (R^2) sesuai dengan persamaan 2.22.
 - g. Melakukan interpretasi model.

3.6 Diagram Alir

Berdasarkan langkah analisis yang telah dijabarkan, secara visual dapat diperjelas menggunakan diagram alir yang disajikan pada Gambar 3.1 sebagai berikut.



Gambar 3.1 Diagram Alir Penelitian

(Halaman ini sengaja dikosongkan)

BAB IV ANALISIS DAN PEMBAHASAN

Pada bab ini akan dibahas mengenai karakteristik dari data angka kecelakaan lalu lintas di Jawa Timur beserta faktor-faktor yang diduga mempengaruhinya. Selanjutnya adalah melakukan pemodelan menggunakan regresi nonparametrik Spline dengan langkah pemilihan titik knot optimal, uji parameter model secara serentak dan parsial, dan uji asumsi residual.

4.1 Karakteristik Angka Kecelakaan Lalu Lintas di Provinsi Jawa Timur

Karakteristik data pada penelitian ini meliputi analisis deskriptif dan analisis hubungan pola data antara variabel respon dengan masing-masing variabel prediktor berdasarkan *scatterplot*.

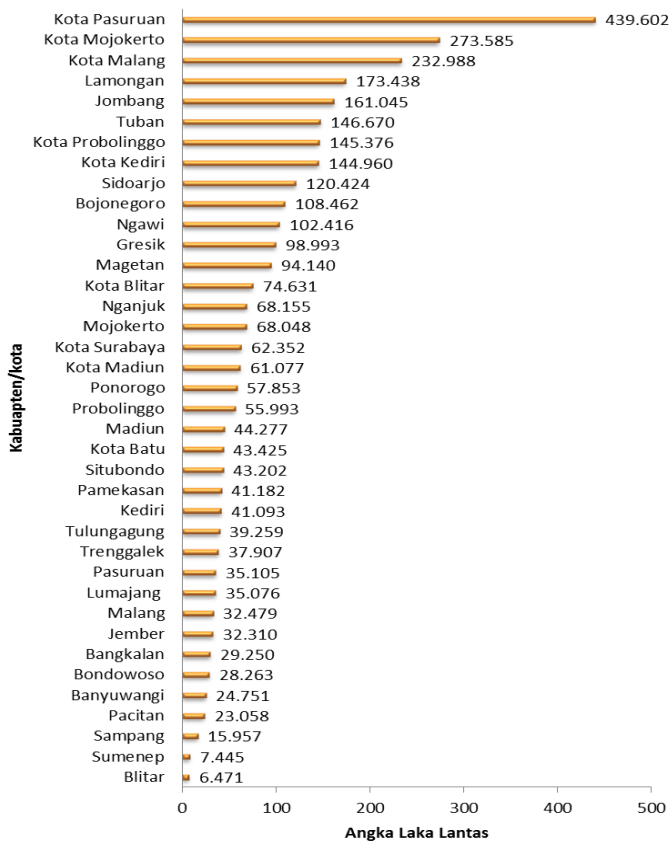
4.1.1 Analisis Deskriptif Angka Kecelakaan Lalu Lintas di Provinsi Jawa Timur

Angka kecelakaan lalu lintas merupakan alat penentu untuk menyatakan tingkat kecelakaan lalu lintas pada satu satuan segmen jalan. Berikut ini adalah karakteristik dari angka kecelakaan lalu lintas di Provinsi Jawa Timur beserta variabel prediktor yang diduga mempengaruhinya menggunakan statistika deskriptif yang disajikan dalam Tabel 4.1.

Tabel 4.1 Statistika Deskriptif Variabel

Variabel	Rata-rata	Standar Deviasi
y	85,545	85,127
x_1	1829,658	2184,106
x_2	8,387	6,065
x_3	75,057	17,615
x_4	14,651	6,720
x_5	57,949	15,015

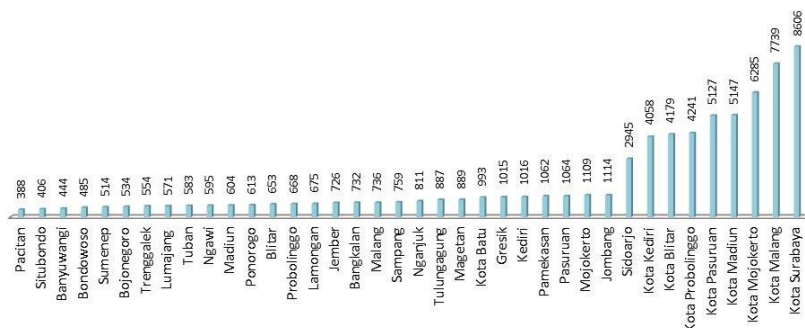
Berdasarkan Tabel 4.1, variabel respon (y) merupakan angka kecelakaan di Provinsi Jawa Timur dengan nilai standar deviasi sebesar 85,127 dan rata-rata sebesar 85,545. Hal ini berarti pada tahun 2015, angka kecelakaan lalu lintas Jawa Timur cukup tinggi yaitu di setiap 100 kilometer rata-rata terjadi 85,545 atau 86 kecelakaan lalu lintas. Terdapat 13 kabupaten/kota di Jawa Timur yang memiliki angka kecelakaan lalu lintas lebih tinggi dari rata-rata dan 25 kabupaten/kota di Jawa Timur yang memiliki angka kecelakaan lalu lintas lebih rendah dari rata-rata.



Gambar 4.1 Angka Kecelakaan Lalu Lintas di Jawa Timur

Pada Gambar 4.1, angka kecelakaan terendah setiap 100 kilometer sebesar 6,471 atau 7 kecelakaan berada di Kabupaten Blitar dan angka kecelakaan tertinggi setiap 100 kilometer sebesar 439,602 atau 440 kecelakaan terdapat di Kota Pasuruan.

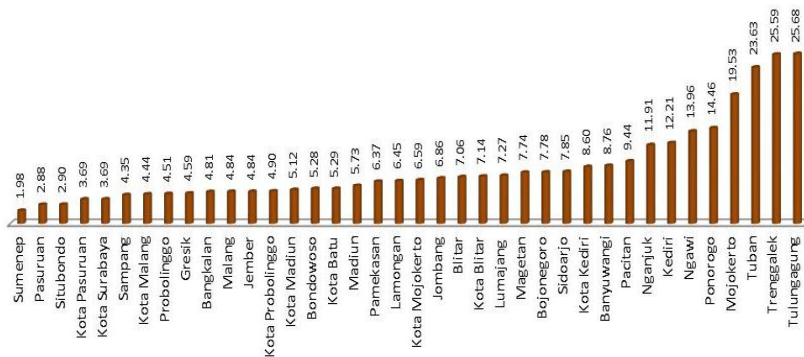
Variabel x_1 merupakan variabel kepadatan penduduk, dimana Jawa Timur merupakan salah satu Provinsi yang memiliki kepadatan penduduk tertinggi di Indonesia, sehingga variabel ini diduga berpengaruh terhadap angka kecelakaan lalu lintas di Jawa Timur. Pada Tabel 4.1 diketahui bahwa variabel kepadatan penduduk memiliki rata-rata sebesar 1829,658 dengan standar deviasi sebesar 2184,106. Hal ini menunjukkan bahwa dalam setiap 1 kilometer persegi, rata-rata daerah di Jawa Timur ditempati oleh 1830 jiwa. Terdapat 9 kabupaten/kota di Jawa Timur yang memiliki kepadatan penduduk di atas 1830 jiwa/km² dan 29 kabupaten/kota yang memiliki kepadatan penduduk kurang dari 1830 jiwa/km². Pada Gambar 4.2 menjelaskan bahwa kepadatan penduduk yang paling rendah sebesar 388 jiwa/km² di Kabupaten Pacitan, sedangkan kepadatan penduduk di Provinsi Jawa Timur yang paling tinggi sebesar 8606 jiwa/km² di Kota Surabaya.



Gambar 4.2 Kepadatan Penduduk di Jawa Timur

Variabel x_2 merupakan variabel persentase pelaku kecelakaan usia remaja, dimana kelompok umur tersebut merupakan penyumbang kecelakaan tertinggi. Oleh karena itu, persentase

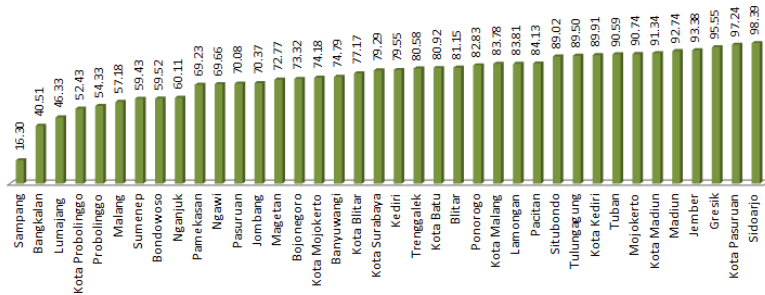
pelaku usia remaja diduga berpengaruh dalam menentukan angka kecelakaan lalu lintas di Jawa Timur. Pada Tabel 4.1 diketahui bahwa variabel persentase pelaku kecelakaan usia remaja memiliki rata-rata sebesar 8,387 persen dengan standar deviasi sebesar 6,065 persen. Terdapat 27 kabupaten/kota yang memiliki persentase pelaku usia remaja di bawah 8,387 persen dan 11 kabupaten/kota yang memiliki persentase usia remaja di atas rata-rata persentase pelaku usia remaja di Jawa Timur. Berdasarkan Gambar 4.3 nilai persentase pelaku usia remaja tertinggi sebesar 25,68 persen di Kabupaten Tulungagung, sedangkan yang terendah berada di Kabupaten Sumenep sebesar 1,98 persen.



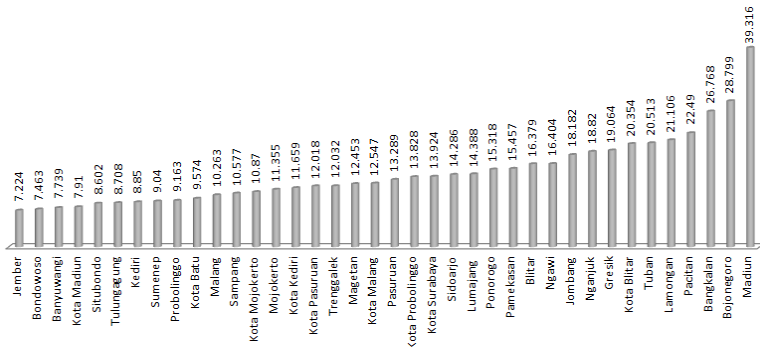
Gambar 4.3 Pelaku Kecelakaan Usia Remaja di Jawa Timur

Variabel x_3 merupakan variabel persentase pelaku kecelakaan berpendidikan terakhir di atas SMP. Perilaku berkendara dipengaruhi oleh tingkat pendidikan terakhir yang dicapai. Wahyuningtyas dalam Dewi (2016) menyatakan bahwa terdapat perbedaan yang signifikan antara pelaku kecelakaan yang pendidikan terakhirnya SD, SMP, SMA atau Perguruan Tinggi. Dalam Tabel 4.1 diketahui bahwa variabel pelaku kecelakaan berpendidikan terakhir di atas SMP memiliki persentase rata-rata sebesar 75,057 persen dengan standar deviasi sebesar 17,615. Terdapat 17 kabupaten/kota yang memiliki persentase pelaku kecelakaan di atas SMP di bawah rata-rata dan 21 kabupaten/kota

yang memiliki persentase pelaku kecelakaan di atas SMP di atas rata-rata. Berdasarkan Gambar 4.4 diperoleh persentase terendah pada pelaku kecelakaan berpendidikan terakhir di atas SMP sebesar 16,30 persen berada di Kabupaten Sampang, sedangkan persentase tertinggi sebesar 98,39 persen terdapat di Kabupaten Sidoarjo.



Gambar 4.4 Pelaku Berpendidikan Akhir Lebih dari SMP di Jawa Timur

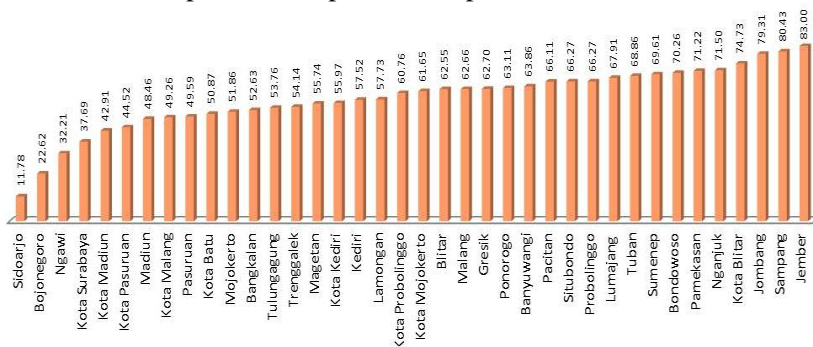


Gambar 4.5 Rasio Jenis Kelamin Pelaku di Jawa Timur

Variabel x_4 merupakan variabel rasio jenis kelamin pelaku yang diduga berpengaruh terhadap angka kecelakaan lalu lintas di Jawa Timur. Berdasarkan Tabel 4.1, diketahui bahwa rata-rata rasio jenis kelamin pelaku per 100 pelaku kecelakaan di Provinsi Jawa Timur sebesar 14,651 yang menunjukkan bahwa rata-rata jenis kelamin pelaku laki-laki lebih banyak daripada perempuan.

Berdasarkan Gambar 4.5 diperoleh informasi bahwa rasio jenis kelamin pelaku terendah berada di Kabupaten Jember yakni sebesar 7,224 dan rasio tertinggi sebesar 39,316 terdapat di Kabupaten Madiun.

Variabel x_5 merupakan variabel persentase pelaku yang tidak memiliki SIM yang diduga berpengaruh terhadap angka kecelakaan lalu lintas di Jawa Timur. Menurut Bakharuddin dalam Satria (2014), salah satu penyebab tingginya angka kecelakaan dikarenakan tingginya jumlah pengemudi yang belum memiliki SIM. Pada Tabel 4.1 variabel persentase pelaku tidak memiliki SIM mempunyai rata-rata sebesar 57,949 persen dengan standar deviasi sebesar 15,015. Hal ini menunjukkan bahwa terdapat 18 kabupaten/kota di Jawa Timur dengan persentase pelaku yang tidak memiliki SIM berada di bawah rata-rata dan terdapat 20 kabupaten/kota yang berada di atas persentase rata-rata pada variabel pelaku yang tidak memiliki SIM. Berdasarkan Gambar 4.6 menginformasikan bahwa nilai terendah pada persentase pelaku yang tidak memiliki SIM sebesar 11,78 persen terdapat di Kabupaten Sidoarjo, sedangkan persentase tertinggi sebesar 83,00 persen terdapat di Kabupaten Jember.



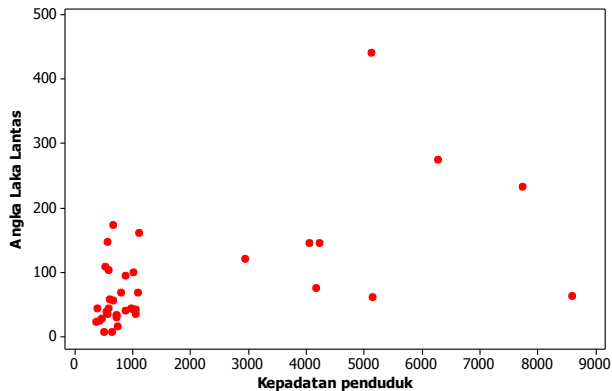
Gambar 4.6 Pelaku Kecelakaan Tidak Memiliki SIM di Jawa Timur

Berdasarkan Tabel 4.1 diketahui bahwa standar deviasi untuk variabel kepadatan penduduk lebih tinggi dibandingkan variabel lainnya. Hal ini menunjukkan bahwa variabel kepadatan

penduduk merupakan variabel yang paling beragam dalam penelitian angka kecelakaan lalu lintas di Provinsi Jawa Timur.

4.1.2 Analisis Pola Hubungan Angka Kecelakaan Lalu Lintas dengan Faktor-Faktor yang Diduga Mempengaruhi

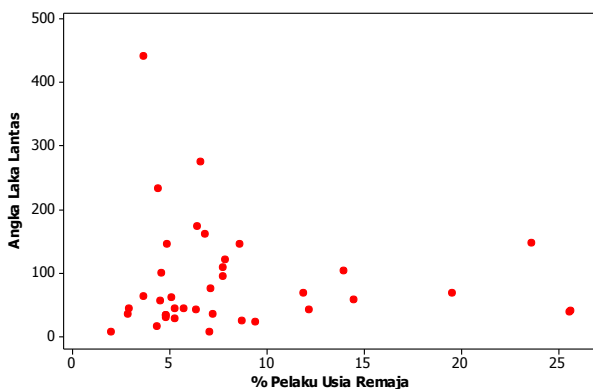
Salah satu analisis karakteristik angka kecelakaan lalu lintas di Jawa Timur adalah dengan membuat *scatterplot* antara variabel respon dengan variabel prediktor untuk mengetahui pola hubungan angka kecelakaan lalu lintas dengan masing-masing variabel prediktor yang diduga mempengaruhinya.



Gambar 4.7 *Scatterplot* Antara Angka Laka dengan Kepadatan Penduduk

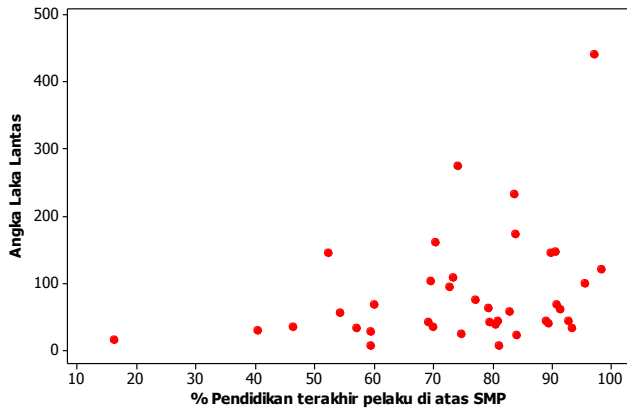
Dapat diketahui dari Gambar 4.7 variabel angka laka lalu lintas dengan kepadatan penduduk sebaran plotnya tidak membentuk pola tertentu, dan juga pola tersebut cenderung berubah pada sub-sub interval tertentu sehingga termasuk dalam komponen nonparametrik. Pada umumnya, semakin padat jumlah penduduk di suatu wilayah maka angka kecelakaan lalu lintas di wilayah tersebut akan cenderung meningkat dikarenakan mobilitas penduduk dari satu tempat ke tempat lain menggunakan fasilitas jalan raya juga cukup tinggi, namun tidak semua wilayah yang padat penduduk mendapati angka kecelakaan lalu lintas yang tinggi.

Dari Gambar 4.8 diperoleh informasi bahwa pola hubungan angka laka lalu lintas dengan persentase pelaku kecelakaan usia remaja tidak ada kecenderungan membentuk pola tertentu, sehingga dapat dikategorikan dalam komponen nonparametrik. Tingginya pelaku kecelakaan yang disebabkan oleh anak pada usia remaja dapat mempengaruhi angka laka di kabupaten/kota di Jawa Timur, mengingat usia tersebut belum cukup umur untuk mengemudi tanpa pengawasan orang dewasa serta adanya pengaruh tingkat emosi yang belum stabil dalam berkendara.

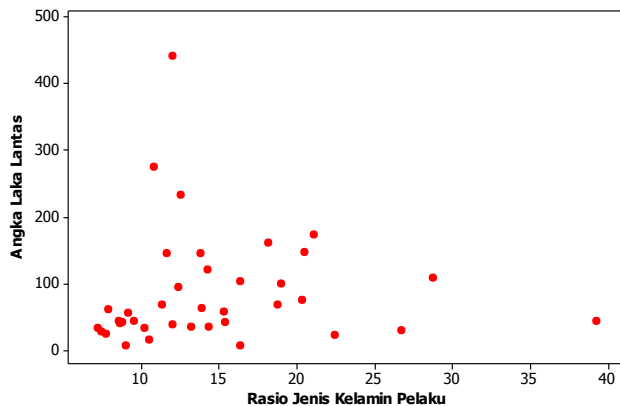


Gambar 4.8 Scatterplot Antara Angka Laka dengan Pelaku Usia Remaja

Gambar 4.9 menunjukkan pola hubungan antara angka laka lalu lintas dengan persentase pelaku kecelakaan berpendidikan terakhir di atas SMP tidak membentuk pola tertentu. Angka laka lintas dapat dipengaruhi oleh pelaku kecelakaan dengan pendidikan terakhir lebih dari SMP. Hal ini disebabkan adanya sikap disiplin berlalu lintas yang berbeda antar jenjang pendidikan yang ditamatkan. Pada umumnya, seseorang yang telah menyelesaikan pendidikan lebih dari SMP dianggap memiliki etika berkendara yang lebih baik, namun tidak semua pelaku kecelakaan yang berpendidikan lebih dari SMP memiliki etika berkendara yang baik dalam berkendara, sehingga variabel ini merupakan komponen nonparametrik.



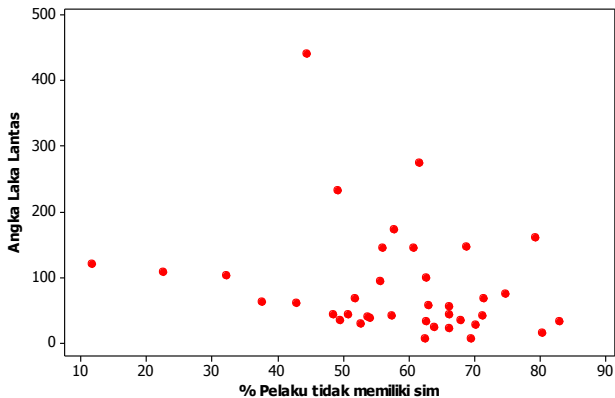
Gambar 4.9 Scatterplot Antara Angka Laka dengan Pelaku Lebih dari SMP



Gambar 4.10 Scatterplot Antara Angka Laka dengan Rasio JK Pelaku

Rasio jenis kelamin pelaku merupakan perbandingan antara pelaku kecelakaan wanita dan pelaku laki-laki per 100 pelaku pada suatu wilayah tertentu. Berdasarkan Gambar 4.10 pola hubungan antara angka laka dan rasio jenis kelamin pelaku tidak membentuk pola tertentu. Apabila angka laka rendah, tidak tentu rasio jenis kelamin pelaku di kabupaten/kota Jawa Timur rendah

dan apabila angka laka tinggi, tidak tentu rasio jenis kelamin pelaku di kabupaten/kota Jawa Timur tinggi, sehingga rasio jenis kelamin pelaku merupakan komponen nonparametrik.



Gambar 4.11 Scatterplot Antara Angka Laka dengan Pelaku Tanpa SIM

Gambar 4.11 menunjukkan pola hubungan antara angka laka lantas dengan persentase pelaku kecelakaan yang tidak memiliki SIM sebaran plotnya membentuk pola yang acak sehingga dapat termasuk dalam komponen nonparametrik. Kondisi ini dikaitkan dengan pengetahuan yang dimiliki oleh pengendara. Pengemudi yang telah memiliki SIM dianggap telah mengerti tata cara berkendara yang baik di jalan raya karena telah melalui proses ujian secara tertulis maupun praktik saat pembuatan SIM berlangsung, sehingga berpengaruh terhadap angka laka lantas di kabupaten/kota di Jawa Timur.

4.2 Pemodelan Angka Kecelakaan Lalu Lintas di Jawa Timur Menggunakan Regresi Nonparametrik Spline

Sebelumnya telah dilakukan analisis deksriptif untuk mengetahui hubungan pola data antara variabel respon dan variabel prediktor yang diduga berpengaruh menggunakan *scatterplot* dan didapatkan hasil seluruh variabel masuk dalam komponen nonparametrik sehingga pemodelan angka kecelakaan

lalu lintas di Jawa Timur dengan faktor-faktor yang diduga memengaruhinya menggunakan metode regresi nonparametrik Spline.

4.2.1 Memodelkan Menggunakan Titik Knot

Model terbaik didapatkan dari titik knot yang optimum. Untuk memperoleh titik knot yang optimum, dapat dilihat dari nilai GCV yang paling minimum. Berikut merupakan pemilihan titik knot optimal dengan satu knot, dua knot, tiga knot, dan kombinasi knot.

a. Satu Titik Knot

Estimasi model regresi nonparametrik Spline dengan satu titik knot pada angka kecelakaan lalu lintas di kabupaten/kota di Provinsi Jawa Timur adalah sebagai berikut.

$$\hat{y} = \hat{\beta}_0 + \hat{\beta}_1 x_1 + \hat{\beta}_2 (x_1 - k_1)_+^1 + \hat{\beta}_3 x_2 + \hat{\beta}_4 (x_2 - k_2)_+^1 + \hat{\beta}_5 x_3 + \hat{\beta}_6 (x_3 - k_3)_+^1 + \hat{\beta}_7 x_4 + \hat{\beta}_8 (x_4 - k_4)_+^1 + \hat{\beta}_9 x_5 + \hat{\beta}_{10} (x_5 - k_5)_+^1$$

Tabel 4.2 menunjukkan sepuluh nilai GCV yang berada di sekitar nilai GCV paling minimum untuk model regresi nonparametrik Spline satu titik knot.

Tabel 4.2 Titik Knot dan GCV untuk Satu Knot

No	Knot					GCV
	x_1	x_2	x_3	x_4	x_5	
1	6928,86	20,84	81,64	32,77	68,47	6313,01
2	7096,57	21,32	83,31	33,42	69,92	6249,90
3	7264,29	21,81	84,99	34,08	71,38	6131,91
4	7432,00	22,29	86,66	34,73	72,83	5977,80
5	7599,71	22,77	88,34	35,39	74,28	5809,16
6	7767,43	23,26	90,01	36,04	75,74	5626,40
7	7935,14	23,74	91,69	36,70	77,19	5359,52
8	8102,86	24,23	93,36	37,35	78,64	5329,14
9	8270,57	24,71	95,04	38,01	80,10	5496,00
10	8438,29	25,19	96,71	38,66	81,55	6258,35

Tabel 4.2 menunjukkan bahwa nilai GCV pada satu titik knot yang paling minimum adalah 5329,14 dengan titik knot optimum yaitu variabel kepadatan penduduk (x_1) berada pada titik knot 8102,86; variabel persentase pelaku usia remaja (x_2) berada pada titik knot 24,23; variabel persentase pendidikan terakhir pelaku adalah di atas SMP (x_3) berada pada titik knot 93,36; variabel rasio jenis kelamin pelaku (x_4) berada pada titik knot 37,35; serta variabel persentase pelaku kecelakaan tidak memiliki SIM (x_5) berada pada titik knot 78,64.

b. Dua Titik Knot

Setelah dilakukan pemilihan titik knot optimum dengan satu titik knot, selanjutnya dilakukan menggunakan dua titik knot pada setiap variabel. Berikut merupakan estimasi model regresi nonparametrik Spline pada angka kecelakaan lalu lintas di Provinsi Jawa Timur menggunakan dua titik knot.

$$\begin{aligned}\hat{y} = & \hat{\beta}_0 + \hat{\beta}_1 x_1 + \hat{\beta}_2 (x_1 - k_1)_+^1 + \hat{\beta}_3 (x_1 - k_2)_+^1 + \hat{\beta}_4 x_2 + \hat{\beta}_5 (x_2 - k_3)_+^1 + \\ & \hat{\beta}_6 (x_2 - k_4)_+^1 + \hat{\beta}_7 x_3 + \hat{\beta}_8 (x_3 - k_5)_+^1 + \hat{\beta}_9 (x_3 - k_6)_+^1 + \hat{\beta}_{10} x_4 + \\ & \hat{\beta}_{11} (x_4 - k_7)_+^1 + \hat{\beta}_{12} (x_4 - k_8)_+^1 + \hat{\beta}_{13} x_5 + \hat{\beta}_{14} (x_5 - k_9)_+^1 + \\ & \hat{\beta}_{15} (x_5 - k_{10})_+^1\end{aligned}$$

Tabel 4.3 menunjukkan sepuluh nilai GCV yang berada di sekitar nilai GCV paling minimum untuk model regresi nonparametrik Spline dua titik knot, dimana nilai GCV minimum yang diperoleh dari pemodelan menggunakan dua titik knot adalah 3791,55. Titik knot optimum untuk variabel kepadatan penduduk (x_1) berada pada titik knot 8270,57 dan 8438,29; variabel persentase pelaku usia remaja (x_2) berada pada titik knot 24,71 dan 25,19; variabel persentase pendidikan terakhir pelaku adalah di atas SMP (x_3) berada pada titik knot 95,04 dan 96,71; variabel rasio jenis kelamin pelaku (x_4) berada pada titik knot 38,01 dan 38,66; serta variabel persentase pelaku tidak memiliki SIM (x_5) berada pada titik knot 80,10 dan 81,55.

Tabel 4.3 Titik Knot dan GCV untuk Dua Knot

No	Knot					GCV
	x_1	x_2	x_3	x_4	x_5	
1	7935,14	23,74	91,69	36,70	77,19	5725,10
	8102,86	24,23	93,36	37,35	78,64	
2	7935,14	23,74	91,69	36,70	77,19	5966,11
	8270,57	24,71	95,04	38,01	80,10	
3	7935,14	23,74	91,69	36,70	77,19	5256,38
	8438,29	25,19	96,71	38,66	81,55	
4	7935,14	23,74	91,69	36,70	77,19	5359,52
	8606,00	25,68	98,39	39,32	83,01	
5	8102,86	24,23	93,36	37,35	78,64	6079,32
	8270,57	24,71	95,04	38,01	80,10	
6	8102,86	24,23	93,36	37,35	78,64	5022,29
	8438,29	25,19	96,71	38,66	81,55	
7	8102,86	24,23	93,36	37,35	78,64	5329,14
	8606,00	25,68	98,39	39,32	83,01	
8	8270,57	24,71	95,04	38,01	80,10	3791,55
	8438,29	25,19	96,71	38,66	81,55	
9	8270,57	24,71	95,04	38,01	80,10	5496,00
	8606,00	25,68	98,39	39,32	83,01	
10	8438,29	25,19	96,71	38,66	81,55	6258,35
	8606,00	25,68	98,39	39,32	83,01	

Nilai GCV dengan dua titik knot bernilai lebih kecil dari satu titik knot, sehingga model dengan dua titik knot lebih baik daripada dengan satu titik knot. Akan tetapi masih harus dicobakan menggunakan tiga titik knot dan kombinasi titik knot untuk mendapatkan kemungkinan model yang lebih baik.

c. Tiga Titik Knot

Selanjutnya dilakukan pemilihan titik knot optimum menggunakan tiga titik knot pada setiap variabel. Berikut merupakan estimasi model regresi nonparametrik Spline pada angka kecelakaan lalu lintas di Provinsi Jawa Timur menggunakan tiga titik knot.

$$\begin{aligned}\hat{y} = & \hat{\beta}_0 + \hat{\beta}_1 x_1 + \hat{\beta}_2 (x_1 - k_1)_+^1 + \hat{\beta}_3 (x_1 - k_2)_+^1 + \hat{\beta}_4 (x_1 - k_3)_+^1 + \\ & \hat{\beta}_5 x_2 + \hat{\beta}_6 (x_2 - k_4)_+^1 + \hat{\beta}_7 (x_2 - k_5)_+^1 + \hat{\beta}_8 (x_2 - k_6)_+^1 + \\ & \hat{\beta}_9 x_3 + \hat{\beta}_{10} (x_3 - k_7)_+^1 + \hat{\beta}_{11} (x_3 - k_8)_+^1 + \hat{\beta}_{12} (x_3 - k_9)_+^1 + \\ & \hat{\beta}_{13} x_4 + \hat{\beta}_{14} (x_4 - k_{10})_+^1 + \hat{\beta}_{15} (x_4 - k_{11})_+^1 + \hat{\beta}_{16} (x_4 - k_{12})_+^1 + \\ & \hat{\beta}_{17} x_5 + \hat{\beta}_{18} (x_5 - k_{13})_+^1 + \hat{\beta}_{19} (x_5 - k_{14})_+^1 + \hat{\beta}_{20} (x_5 - k_{15})_+^1\end{aligned}$$

Berikut ini merupakan sepuluh nilai GCV minimum untuk model regresi nonparametrik Spline menggunakan tiga titik knot yang disajikan dalam Tabel 4.4.

Tabel 4.4 Titik Knot dan GCV untuk Tiga Knot

No	Knot					GCV
	x_1	x_2	x_3	x_4	x_5	
1	7767,43	23,26	90,01	36,04	75,74	5536,01
	7935,14	23,74	91,69	36,70	77,19	
	8102,86	24,23	93,36	37,35	78,64	
2	7767,43	23,26	90,01	36,04	75,74	6015,33
	7935,14	23,74	91,69	36,70	77,19	
	8270,57	24,71	95,04	38,01	80,10	
3	7767,43	23,26	90,01	36,04	75,74	3894,53
	7935,14	23,74	91,69	36,70	77,19	
	8438,29	25,19	96,71	38,66	81,55	
4	7767,43	23,26	90,01	36,04	75,74	6028,38
	8102,86	24,23	93,36	37,35	78,64	
	8270,57	24,71	95,04	38,01	80,10	

Tabel 4.4 Titik Knot dan GCV untuk Tiga Knot (Lanjutan)

No	Knot					GCV
	x_1	x_2	x_3	x_4	x_5	
5	7767,43	23,26	90,01	36,04	75,74	2794,34
	8102,86	24,23	93,36	37,35	78,64	
	8438,29	25,19	96,71	38,66	81,55	
6	7767,43	23,26	90,01	36,04	75,74	3203,26
	8270,57	24,71	95,04	38,01	80,10	
	8438,29	25,19	96,71	38,66	81,55	
7	7935,14	23,74	91,69	36,70	77,19	6131,32
	8102,86	24,23	93,36	37,35	78,64	
	8270,57	24,71	95,04	38,01	80,10	
8	7935,14	23,74	91,69	36,70	77,19	3368,18
	8102,86	24,23	93,36	37,35	78,64	
	8438,29	25,19	96,71	38,66	81,55	
9	7935,14	23,74	91,69	36,70	77,19	3513,06
	8270,57	24,71	95,04	38,01	80,10	
	8438,29	25,19	96,71	38,66	81,55	
10	8102,86	24,23	93,36	37,35	78,64	3602,30
	8270,57	24,71	95,04	38,01	80,10	
	8438,29	25,19	96,71	38,66	81,55	

Tabel 4.4 menginformasikan bahwa nilai GCV minimum yang diperoleh dari pemodelan menggunakan tiga titik knot adalah 2794,34. Titik knot optimum untuk variabel kepadatan penduduk (x_1) berada pada titik knot 7767,43, 8102,86, dan 8438,29; variabel persentase pelaku usia remaja (x_2) berada pada titik knot 23,26, 24,23, dan 25,19; variabel persentase pendidikan terakhir pelaku adalah di atas SMP (x_3) berada pada titik knot 90,01, 93,36, dan 96,71; variabel rasio jenis kelamin pelaku (x_4) berada pada titik knot 36,04, 37,35, dan 38,66; serta variabel

persentase pelaku kecelakaan tidak memiliki SIM (x_5) berada pada titik knot 75,74, 78,64 dan 81,55.

Nilai GCV dengan tiga titik knot memiliki nilai yang lebih kecil dari satu titik knot dan dua titik knot, sehingga model dengan tiga titik knot dapat dikatakan lebih baik. Namun, masih harus dicobakan menggunakan kombinasi titik knot untuk mendapatkan kemungkinan model yang lebih baik.

d. Kombinasi Titik Knot

Selanjutnya adalah pemilihan titik knot optimum dengan kombinasi titik knot. Hal ini dilakukan karena terdapat kemungkinan jumlah titik knot optimum dari setiap variabel prediktor berbeda-beda. Berikut merupakan sepuluh nilai GCV yang berada disekitar nilai GCV minimum yang disajikan dalam Tabel 4.5.

Tabel 4.5 Titik Knot dan GCV untuk Kombinasi Knot

No	Knot					GCV
	x_1	x_2	x_3	x_4	x_5	
1	8102,86	23,26	90,01	38,01	80,09	3579,19
		24,23	93,36	38,66	81,55	
		25,19	96,73			
2	8102,86	23,26	90,01	38,01	75,74	2794,36
		24,23	93,36	38,66	78,64	
		25,19	96,73		81,55	
3	8102,86	23,26	90,01	36,04	78,64	3386,50
		24,23	93,36	37,35		
		25,19	96,73	38,66		
4	8102,86	23,26	90,01	36,04	80,09	3579,17
		24,23	93,36	37,35	81,55	
		25,19	96,73	38,66		
5	8102,86	23,26	90,01	36,04	75,74	2794,34
		24,23	93,36	37,35	78,64	
		25,19	96,73	38,66	81,55	
6	8270,57	24,23	93,36	37,35	78,64	5329,14
	8438,29					

Tabel 4.2 Titik Knot dan GCV untuk Kombinasi Knot (Lanjutan)

No	Knot					GCV
	x_1	x_2	x_3	x_4	x_5	
7	8270,57	24,23	93,36	37,35	80,09	5716,68
	8438,29				81,55	
	8270,57					
8	8438,29	24,23	93,36	37,35	75,74	5809,84
	8270,57				78,64	
	8438,29				81,55	
9	8438,29	24,23	93,36	38,01	78,64	5329,14
	8270,57			38,66		
	8438,29					
10	8438,29	24,23	93,36	38,01	80,09	5716,68
	8270,57			38,66	81,55	
	8438,29					

Berdasarkan Tabel 4.5 diketahui bahwa nilai GCV minimum dari kombinasi knot adalah 2794,3355. Nilai GCV tersebut dihasilkan pada kombinasi titik knot 1,3,3,3,3. Titik knot optimum untuk variabel kepadatan penduduk (x_1) berada pada titik knot 8102,86; variabel persentase pelaku usia remaja (x_2) berada pada titik knot 23,26, 24,23, dan 25,19; variabel persentase pendidikan terakhir pelaku adalah di atas SMP (x_3) berada pada titik knot 90,01, 93,36, dan 96,73; variabel rasio jenis kelamin pelaku (x_4) berada pada titik knot 36,04, 37,35, dan 38,66; serta variabel persentase pelaku tidak memiliki SIM (x_5) berada pada titik knot 75,74, 78,64, dan 81,55.

4.2.2 Pemilihan Titik Knot Optimal

Titik knot optimal merupakan titik knot yang mempunyai nilai GCV paling minimum. Perbandingan nilai GCV minimum yang diperoleh dari satu titik knot, dua titik knot, tiga titik knot, dan kombinasi knot yang disajikan pada Tabel 4.6 dapat diperoleh informasi bahwa nilai GCV minimum pada tiga titik knot dan kombinasi knot adalah sama. Dalam penelitian ini dipilih titik knot yang memiliki jumlah parameter paling sedikit sesuai dengan teori *parsimony*, dimana suatu model yang baik

adalah model yang memiliki jumlah parameter paling sederhana. Oleh karena itu, diputuskan titik knot optimum yang terpilih adalah kombinasi titik knot (1,3,3,3,3).

Tabel 4.6 Perbandingan Nilai GCV Minimum

Jumlah Knot	GCV Minimum
Satu Titik Knot	5329,13551947391
Dua Titik Knot	3791,55173130658
Tiga Titik Knot	2794,3354549638
Kombinasi Knot (1,3,3,3,3)	2794,33545685926

4.2.3 Model Terbaik untuk Angka Kecelakaan Lalu Lintas di Provinsi Jawa Timur

Berdasarkan kriteria pemilihan model terbaik, digunakan titik knot optimum yang telah terpilih berdasarkan nilai GCV paling minimum yaitu menggunakan kombinasi knot (1,3,3,3,3). Berikut merupakan model regresi nonparametrik Spline terbaik.

$$\begin{aligned} \hat{y} = & 93,015 + 0,029x_1 - 0,488(x_1 - 8102,86)_+^1 + 3,566x_2 + 316,341(x_2 - 23,26)_+^1 - \\ & 532,815(x_2 - 24,23)_+^1 - 211,674(x_2 - 25,19)_+^1 + 0,177x_3 - 118,672(x_3 - 90,01)_+^1 + \\ & 338,428(x_3 - 93,36)_+^1 - 463,59(x_3 - 96,71)_+^1 - 0,41x_4 + 68,385(x_4 - 36,04)_+^1 + \\ & 41,031(x_4 - 37,35)_+^1 + 13,677(x_4 - 38,66)_+^1 - 1,589x_5 + 69,061(x_5 - 75,74)_+^1 - \\ & 173,13(x_5 - 78,64)_+^1 + 458,183(x_5 - 81,55)_+^1 \end{aligned}$$

4.2.4 Pengujian Signifikansi Parameter

Pengujian parameter dilakukan untuk mengetahui apakah parameter yang telah didapatkan dari hasil pemodelan dengan regresi nonparametrik Spline signifikan atau tidak terhadap variabel respon. Pengujian dilakukan dengan dua tahap, yaitu pengujian parameter secara serentak dan parsial.

a. Pengujian Serentak

Pengujian secara serentak dilakukan pada parameter model regresi terhadap variabel respon secara bersama-sama atau serentak. Hipotesis yang digunakan untuk pengujian secara serentak adalah sebagai berikut.

$$H_0 : \beta_1 = \beta_2 = \dots = \beta_{18} = 0$$

$$H_1 : \text{minimal terdapat satu } \beta_g \neq 0, g = 1, 2, \dots, 18$$

Hasil pengujian secara serentak ditampilkan pada Tabel 4.7.

Tabel 4.7 Hasil Pengujian Secara Serentak

Sumber	df	SS	MS	F
Regresi	18	232534,90	12918,61	6,897
Error	19	35590,87	1873,20	
Total	37	268125,80		

Berdasarkan Tabel 4.7, dengan menggunakan taraf signifikan sebesar 5% maka didapatkan keputusan tolak H_0 karena nilai $F_{hitung} = 5,863 > F_{(0,05;18,19)} = 2,182$. Hal ini menunjukkan bahwa minimal terdapat satu parameter yang signifikan pada model atau secara bersama terdapat minimal satu variabel prediktor berpengaruh secara signifikan terhadap angka kecelakaan lalu lintas di Jawa Timur.

b. Pengujian Parsial

Berdasarkan hasil pengujian serentak, maka dilanjutkan pada pengujian parameter secara parsial. Variabel prediktor dikatakan berpengaruh terhadap variabel respon apabila terdapat minimal satu parameter yang signifikan. Adapun pengujian dilakukan dengan hipotesis sebagai berikut.

$$H_0 : \beta_g = 0$$

$$H_1 : \beta_g \neq 0, g = 1, 2, \dots, 18$$

Pada Tabel 4.8 diketahui bahwa kelima variabel prediktor mempunyai minimal satu parameter yang signifikan terhadap model karena memiliki nilai $|t_{hitung}| > t_{(0,025;19)} = 2,093$, sehingga variabel kepadatan penduduk (x_1), variabel persentase pelaku usia remaja (x_2), variabel persentase pendidikan terakhir pelaku adalah di atas SMP (x_3), variabel rasio jenis kelamin pelaku (x_4), dan variabel persentase pelaku kecelakaan tidak memiliki SIM (x_5) berpengaruh terhadap angka kecelakaan lalu lintas di Jawa Timur.

Tabel 4.8 Hasil Pengujian Secara Parsial

Variabel	Parameter	Estimasi	<i>t</i>	Keputusan	Kesimpulan
Konstan	β_0	93,015	1,135	Gagal Tolak	Tidak sig.
x_1	β_1	0,029	6,723	Tolak	x_1 Berpengaruh
	β_2	-0,488	-4,533	Tolak	
x_2	β_3	3,566	1,504	Gagal Tolak	x_2 Berpengaruh
	β_4	316,341	2,048	Gagal Tolak	
	β_5	-532,815	-2,405	Tolak	
	β_6	-211,674	-2,464	Tolak	
x_3	β_7	0,177	0,252	Gagal Tolak	x_3 Berpengaruh
	β_8	-118,672	-4,151	Tolak	
	β_9	338,428	5,203	Tolak	
	β_{10}	-463,959	-5,502	Tolak	
x_4	β_{11}	-0,410	-0,257	Gagal Tolak	x_4 Berpengaruh
	β_{12}	68,385	3,164	Tolak	
	β_{13}	41,031	3,164	Tolak	
	β_{14}	13,677	3,164	Tolak	
x_5	β_{15}	-1,589	-2,037	Gagal Tolak	x_5 Berpengaruh
	β_{16}	69,061	2,467	Tolak	
	β_{17}	-173,130	-1,914	Gagal Tolak	
	β_{18}	458,183	2,853	Tolak	

4.2.5 Pengujian Asumsi Residual

Pengujian asumsi residual dilakukan untuk mengetahui apakah residual yang dihasilkan dari model telah memenuhi asumsi identik, independen, dan berdistribusi normal (IIDN) atau tidak. Apabila suatu model regresi tidak memenuhi asumsi residual IIDN, namun memiliki kriteria model terbaik dan parameter signifikan maka model tersebut tetap tidak layak digunakan untuk memodelkan angka kecelakaan lalu lintas di Jawa Timur.

a. Asumsi Identik

Pengujian asumsi residual identik digunakan untuk mengetahui apakah terjadi kasus heteroskedastisitas atau tidak, dengan kata lain variansi residual dari model harus homogen. Pengujian asumsi residual identik dilakukan menggunakan uji Glejser dengan hipotesis sebagai berikut.

$$H_0 : \sigma_1^2 = \sigma_2^2 = \dots = \sigma_n^2 = \sigma^2$$

$$H_1 : \text{minimal terdapat satu } \sigma_i^2 \neq \sigma^2, i = 1, 2, \dots, n$$

Hasil uji Glejser disajikan dalam Tabel 4.9 sebagai berikut.

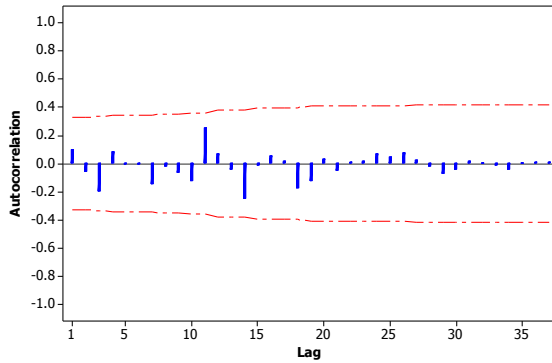
Tabel 4.9 Hasil Pengujian Glejser

Sumber	Df	SS	MS	F
Regresi	5	139,148	279,230	0,490
Error	32	18244,801	570,150	
Total	37	19640,949		

Berdasarkan Tabel 4.9, dengan menggunakan taraf signifikan sebesar 5% maka diperoleh keputusan gagal tolak H_0 karena nilai $F_{hitung} = 0,490 < F_{(0,05;5,32)} = 2,512$. Sehingga dapat disimpulkan bahwa tidak terjadi kasus heteroskedastisitas atau dengan kata lain variansi antar residual homogen. Hal ini berarti asumsi residual identik telah terpenuhi.

b. Asumsi Independen

Residual yang independen adalah residual yang tidak terjadi autokorelasi antar residualnya. Pemeriksaan asumsi independen pada penelitian ini menggunakan plot *Autocorrelation Function* (ACF). Apabila ada nilai autokorelasi yang keluar dari batas atas maupun batas bawah interval konfidensi, maka dapat disimpulkan terdapat kasus autokorelasi antar residual. Gambar 4.12 menunjukkan bahwa tidak terlihat adanya nilai autokorelasi yang keluar batas interval konfidensi (garis merah), sehingga dapat disimpulkan bahwa tidak terdapat autokorelasi antar residual. Hal ini berarti asumsi residual independen telah terpenuhi.



Gambar 4.12 Plot ACF Residual

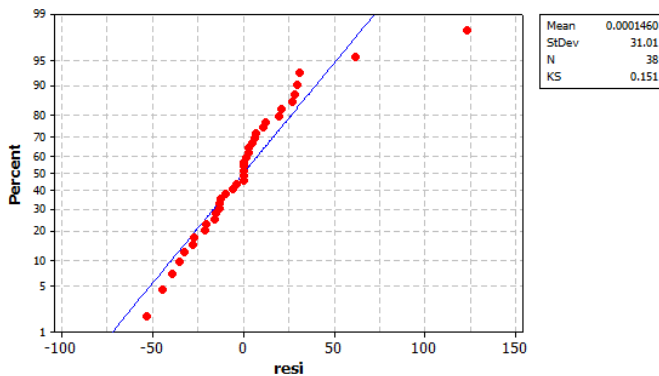
c. Asumsi Distribusi Normal

Pengujian asumsi residual berdistribusi normal dilakukan untuk mengetahui apakah residual telah berdistribusi normal atau tidak. Pengujian ini dilakukan menggunakan uji *Kolmogorov Smirnov* dengan hipotesis sebagai berikut.

$$H_0 : F(x) = F_0(x)$$

$$H_1 : F(x) \neq F_0(x)$$

Hasil uji *Kolmogorov Smirnov* disajikan pada Gambar 4.13 sebagai berikut.



Gambar 4.13 Hasil Uji *Kolmogorov Smirnov*

Berdasarkan Gambar 4.13, dengan menggunakan taraf signifikan sebesar 5% diperoleh nilai $KS = 0,151 < KS_{\text{tabel}} = 0,215$. Sehingga didapatkan keputusan gagal tolak H_0 . Hal ini dapat disimpulkan bahwa residual model telah memenuhi asumsi distribusi normal.

4.2.6 Koefisien Determinasi

Nilai koefisien determinasi (R^2) menunjukkan seberapa besar kebaikan model regresi dalam menjelaskan variabilitas angka kecelakaan lalu lintas di Jawa Timur. Berdasarkan perhitungan didapatkan nilai R^2 sebesar 86,73 persen. Hal ini berarti model regresi nonparametrik Spline yang didapatkan mampu menjelaskan variabilitas angka kecelakaan lalu lintas di Jawa Timur sebesar 86,73 persen, sedangkan sisanya dijelaskan oleh variabel lain yang tidak masuk dalam model.

4.2.7 Interpretasi Model Regresi Nonparametrik Spline

Setelah melakukan pengujian parameter dan semua asumsi residual terpenuhi, selanjutnya adalah menginterpretasikan model regresi yang telah didapatkan. Model yang terbentuk menggunakan titik knot optimum ditunjukkan pada persamaan sebagai berikut.

$$\begin{aligned} \hat{y} = & 93,015 + 0,029x_1 - 0,488(x_1 - 8102,86)_+^1 + 3,566x_2 + 316,341(x_2 - 23,26)_+^1 - \\ & 532,815(x_2 - 24,23)_+^1 - 211,674(x_2 - 25,19)_+^1 + 0,177x_3 - 118,672(x_3 - 90,01)_+^1 + \\ & 338,428(x_3 - 93,36)_+^1 - 463,59(x_3 - 96,71)_+^1 - 0,41x_4 + 68,385(x_4 - 36,04)_+^1 + \\ & 41,031(x_4 - 37,35)_+^1 + 13,677(x_4 - 38,66)_+^1 - 1,589x_5 + 69,061(x_5 - 75,74)_+^1 - \\ & 173,13(x_5 - 78,64)_+^1 + 458,183(x_5 - 81,55)_+^1 \end{aligned}$$

Berdasarkan model tersebut, maka dapat diinterpretasikan dari masing-masing variabel yang berpengaruh terhadap angka kecelakaan lalu lintas di Jawa Timur sebagai berikut.

1. Dengan mengasumsikan variabel x_2 , x_3 , x_4 , dan x_5 konstan, maka pengaruh variabel kepadatan penduduk (x_1) terhadap angka kecelakaan lalu lintas di Jawa Timur dapat ditulis sebagai berikut.

$$\hat{y} = 0,029x_1 - 0,488(x_1 - 8102,86)_+^1$$

$$\hat{y} = \begin{cases} 0,029x_1 & ; x_1 < 8102,86 \\ -0,459x_1 + 3954,196 & ; x_1 \geq 8102,86 \end{cases}$$

Berdasarkan model tersebut, pada interval pertama yaitu kepadatan penduduk kurang dari 8102,86 jiwa/km², apabila kepadatan penduduk mengalami kenaikan sebesar 1 jiwa/km² maka angka kecelakaan lalu lintas di Jawa Timur akan meningkat sebesar 0,029 kecelakaan per 100 kilometer jalan. Terdapat 37 kabupaten/kota yang masuk dalam interval ini yaitu Kabupaten Sumenep, Kabupaten Pasuruan, Kabupaten Situbondo, Kota Pasuruan, Kota Malang, Kabupaten Probolinggo, Kabupaten Gresik, Kabupaten Bangkalan, Kota Probolinggo, Kota Madiun, Kota Batu, Kabupaten Madiun, Kabupaten Sampang, Kabupaten Pamekasan, Kabupaten Lamongan, Kota Mojokerto, Kabupaten Jombang, Kabupaten Blitar, Kota Blitar, Kabupaten Lumajang, Kabupaten Magetan, Kabupaten Bojonegoro, Kabupaten Sidoarjo, Kota Kediri, Kabupaten Banyuwangi, Kabupaten Pacitan, Kabupaten Nganjuk, Kabupaten Kediri, Kabupaten Ngawi, Kabupaten Bondowoso, Kabupaten Ponorogo, Kabupaten Mojokerto, Kabupaten Tuban, Kabupaten Trenggalek, Kabupaten Tulungagung, Kabupaten Malang, dan Kabupaten Jember.

Selanjutnya, kepadatan penduduk lebih dari atau sama dengan 8102,86 jiwa/km², apabila kepadatan penduduk mengalami kenaikan sebesar satu jiwa/km² maka angka kecelakaan lalu lintas di Jawa Timur akan menurun sebesar 0,459 kecelakaan per 100 kilometer jalan. Kabupaten/kota yang termasuk dalam interval ini hanya Kota Surabaya. Pengelompokkan kabupaten/kota secara visual berdasarkan variabel kepadatan penduduk yang mempengaruhi angka kecelakaan lalu lintas di Jawa Timur disajikan pada Gambar 4.14 sebagai berikut.



Gambar 4.14 Persebaran Kab./Kota Berdasarkan Kepadatan Penduduk

2. Dengan mengasumsikan variabel x_1 , x_3 , x_4 , dan x_5 konstan, maka pengaruh variabel persentase pelaku usia remaja (x_2) terhadap angka kecelakaan lalu lintas di Jawa Timur dapat ditulis sebagai berikut.

$$\hat{y} = 3,566x_2 + 316,341(x_2 - 23,26)_+^1 - 532,815(x_2 - 24,23)_+^1 - 211,674(x_2 - 25,19)_+^1$$

$$\hat{y} = \begin{cases} 3,566x_2 & ; x_2 < 23,26 \\ 319,907x_2 - 7358,09 & ; 23,26 \leq x_2 < 24,23 \\ -212,908x_2 + 5552,02 & ; 24,23 \leq x_2 < 25,19 \\ -424,582x_2 - 10884,08 & ; x_2 \geq 25,19 \end{cases}$$

Pada interval pertama yaitu persentase pelaku usia remaja yang kurang dari 23,26 persen, apabila terjadi kenaikan pada persentase pelaku usia remaja sebesar satu persen, maka akan terjadi kenaikan pada angka kecelakaan lalu lintas sebesar 3,566 kecelakaan per 100 kilometer jalan. Terdapat 35 kabupaten/kota di Jawa Timur yang termasuk dalam interval ini yaitu Kabupaten Sumenep, Kabupaten Pasuruan, Kabupaten Situbondo, Kota Pasuruan, Kota Surabaya, Kabupaten Sampang, Kota Malang, Probolinggo, Gresik, Bangkalan, Malang, Jember, Kota Probolinggo, Kota Madiun, Kabupaten Bondowoso, Kota Batu, Kabupaten Madiun,

Kabupaten Pamekasan, Kabupaten Lamongan, Kota Mojokerto, Kabupaten Jombang, Kabupaten Blitar, Kota Blitar, Kabupaten Lumajang, Kabupaten Magetan, Kabupaten Bojonegoro, Kabupaten Sidoarjo, Kota Kediri, Kabupaten Banyuwangi, Kabupaten Pacitan, Kabupaten Nganjuk, Kabupaten Kediri, Kabupaten Ngawi, Kabupaten Ponorogo, dan Kabupaten Mojokerto.

Untuk interval kedua yaitu persentase pelaku usia remaja yang berada di antara 23,26 dan 24,23, apabila terjadi kenaikan sebesar satu persen pada persentase pelaku usia remaja, maka akan terjadi peningkatan angka kecelakaan lalu lintas sebesar 319,907 kecelakaan per 100 kilometer jalan. Kabupaten/kota yang masuk dalam interval ini hanya Kabupaten Tuban.

Kemudian untuk pelaku usia remaja yang berada lebih dari 25,19 persen, apabila terjadi kenaikan persentase pelaku usia remaja sebesar satu persen, maka akan terjadi penurunan pada angka kecelakaan lalu lintas di Jawa Timur sebesar 185,64 kecelakaan per 100 kilometer jalan. Kabupaten/kota yang masuk dalam interval ini adalah Kabupaten Trenggalek dan Kabupaten Tulungagung. Secara visual, pengelompokkan kabupaten/kota berdasarkan variabel persentase pelaku usia remaja disajikan pada Gambar 4.15 sebagai berikut.



Gambar 4.15 Persebaran Kab./Kota Berdasarkan Pelaku Usia Remaja

3. Dengan mengasumsikan variabel x_1 , x_2 , x_4 , dan x_5 konstan, maka pengaruh variabel persentase pelaku berpendidikan terakhir di atas SMP (x_3) terhadap angka kecelakaan lalu lintas di Jawa Timur dapat ditulis sebagai berikut.

$$\hat{y} = 0,177x_3 - 118,672(x_3 - 90,01)_+^1 + 338,428(x_3 - 93,36)_+^1 - 463,59(x_3 - 96,71)_+^1$$

$$\hat{y} = \begin{cases} 0,177x_3 & ; x_3 < 90,01 \\ -118,495x_3 + 10681,667 & ; 90,01 \leq x_3 < 93,36 \\ 219,933x_3 - 20913,971 & ; 93,36 \leq x_3 < 96,71 \\ -244,026x_3 - 23955,504 & ; x_3 \geq 96,71 \end{cases}$$

Untuk interval pertama, apabila kabupaten/kota di Jawa Timur dengan persentase pelaku berpendidikan terakhir di atas SMP kurang dari 90,01 persen mengalami kenaikan sebesar satu persen, maka angka kecelakaan lalu lintas akan meningkat sebesar 0,177 kecelakaan per 100 kilometer. Kabupaten/kota yang tergolong dalam interval ini antara lain Kabupaten Sumenep, Kabupaten Pasuruan, Kabupaten Situbondo, Kota Surabaya, Kabupaten Sampang, Kota Malang, Kabupaten Probolinggo, Kabupaten Bangkalan, Kabupaten Malang, Kota Probolinggo, Kabupaten Pamekasan, Kota Mojokerto, Kabupaten Jombang, Kabupaten Blitar, Kabupaten Lumajang, Kabupaten Bojonegoro, Kota Kediri, Kabupaten Banyuwangi, Kabupaten Pacitan, Kabupaten Nganjuk, Kabupaten Kediri, Kabupaten Ngawi, Kabupaten Trenggalek, Kabupaten Tulungagung, Kota Blitar, Kota Batu, Kabupaten Lamongan, Kabupaten Magetan, Kabupaten Ponorogo, dan Kabupaten Bondowoso. Tanda pada persamaan berbeda dengan teori yang ada. Hal ini bisa saja terjadi karena pelaku kecelakaan yang berpendidikan terakhir lebih dari SMP tidak menjamin angka kecelakaan lalu lintas yang rendah di suatu daerah, serta terdapat faktor lain yang mempengaruhi. Misal di Kabupaten Situbondo, berdasarkan data kecelakaan lalu lintas tahun 2015

yang dicatat oleh Polda Jatim khususnya divisi laka lantas, kejadian kecelakaan berdasarkan faktor pandangan terhalang dan kurangnya penerangan lampu tercatat sebanyak 175 kejadian, jumlah ini tergolong cukup tinggi untuk kasus kecelakaan lalu lintas pada faktor jalan. Setinggi apapun pendidikan yang ditamatkan oleh seseorang, apabila akses jalan untuk penerangannya tidak memadai yang menyebabkan pandangan terhalang, maka kejadian kecelakaan tetap bisa terjadi.

Untuk interval kedua, apabila kabupaten/kota di Jawa Timur dengan persentase pelaku berpendidikan terakhir di atas SMP berada di antara 90,01 dan 93,36 persen naik sebesar satu persen, maka angka kecelakaan lalu lintas akan menurun sebesar 118,495 kecelakaan per 100 kilometer jalan. Terdapat empat kabupaten/kota di Jawa Timur yang masuk dalam interval ini antara lain Kabupaten Tuban, Kabupaten Mojokerto, Kota Madiun, dan Kabupaten Madiun.

Untuk interval ketiga, apabila persentase pelaku berpendidikan terakhir di atas SMP yang berada di antara 93,36 dan 96,71 persen naik sebesar satu persen, maka angka kecelakaan lalu lintas akan meningkat sebesar 219,933 kejadian kecelakaan per 100 kilometer jalan. Terdapat dua kabupaten yang masuk dalam interval ini yaitu Kabupaten Jember dan Kabupaten Gresik.

Untuk interval keempat, apabila kabupaten/kota dengan persentase pelaku berpendidikan terakhir di atas SMP lebih dari 96,71 persen naik sebesar satu persen, maka angka kecelakaan lalu lintas akan menurun sebesar 244,026 kejadian kecelakaan per 100 kilometer jalan. Kabupaten/kota yang masuk dalam interval ini adalah Kota Pasuruan dan Kabupaten Sidoarjo. Secara visual, pengelompokan kabupaten/kota di Jawa Timur berdasarkan variabel persentase pelaku kecelakaan berpendidikan terakhir lebih dari SMP disajikan pada Gambar 4.16 sebagai berikut.



Gambar 4.16 Persebaran Kab./Kota Berdasarkan Pelaku Berpendidikan Akhir Lebih dari SMP

4. Dengan menganggap variabel x_1 , x_2 , x_3 , dan x_5 konstan, maka pengaruh variabel rasio jenis kelamin pelaku (x_4) terhadap angka kecelakaan lalu lintas di Jawa Timur dapat ditulis sebagai berikut.

$$\hat{y} = -0,41x_4 + 68,385(x_4 - 36,04)_+^1 + 41,031(x_4 - 37,35)_+^1 + 13,677(x_4 - 38,66)_+^1$$

$$\hat{y} = \begin{cases} -0,410x_4 & ; x_4 < 36,04 \\ 67,975x_4 - 2464,595 & ; 36,04 \leq x_4 < 37,35 \\ 109,006x_4 - 3997,103 & ; 37,35 \leq x_4 < 38,66 \\ 122,683x_4 - 4525,856 & ; x_4 \geq 38,66 \end{cases}$$

Berdasarkan model yang diperoleh maka dapat diinterpretasikan bahwa apabila kabupaten/kota di Provinsi Jawa Timur dengan rasio jenis kelamin pelaku kurang dari 36,04 dan mengalami pertambahan sebesar satu satuan, maka angka kecelakaan lalu lintas akan turun sebesar 0,41 kejadian kecelakaan per 100 kilometer jalan. Terdapat 37 kabupaten/kota yang termasuk dalam interval ini antara lain Kabupaten Sidoarjo, Kabupaten Bojonegoro, Kabupaten Ngawi, Kabupaten Sumenep, Kabupaten Pasuruan, Kabupaten

Situbondo, Kabupaten Sampang, Kota Malang, Kabupaten Probolinggo, Kabupaten Bangkalan, Kabupaten Malang, Kota Probolinggo, Kabupaten Pamekasan, Kota Mojokerto, Kabupaten Jombang, Kabupaten Blitar, Kabupaten Lumajang, Kota Kediri, Kabupaten Banyuwangi, Kabupaten Pacitan, Kabupaten Nganjuk, Kabupaten Kediri, Kabupaten Trenggalek, Kabupaten Tulungagung, Kota Blitar, Kota Batu, Kabupaten Lamongan, Kabupaten Magetan, Kabupaten Ponorogo, Kabupaten Bondowoso, Kota Pasuruan, Kabupaten Tuban, Kabupaten Mojokerto, Kota Madiun, Kabupaten Jember, Kabupaten Gresik, dan Kota Surabaya.

Kemudian jika rasio jenis kelamin pelaku yang berada diatas 100,45 mengalami kenaikan sebesar satu satuan, maka angka kecelakaan lalu lintas akan naik sebesar 122,683 kejadian kecelakaan per 100 kilometer jalan. Kabupaten yang masuk dalam interval ini adalah hanya Kabupaten Madiun. Pengelompokan kabupaten/kota di Jawa Timur berdasarkan variabel rasio jenis kelamin pelaku secara visual disajikan pada Gambar 4.17.



Gambar 4.17 Persebaran Kab./Kota Berdasarkan Rasio JK Pelaku

5. Dengan menganggap variabel x_1 , x_2 , x_3 , dan x_4 konstan, maka pengaruh variabel persentase pelaku tidak memiliki SIM (x_5)

terhadap angka kecelakaan lalu lintas di Jawa Timur dapat ditulis sebagai berikut.

$$\hat{y} = -1,589x_5 + 69,061(x_5 - 75,74)_+^1 - 173,13(x_5 - 78,64)_+^1 + 458,183(x_5 - 81,55)_+^1$$

$$\hat{y} = \begin{cases} -1,589x_5 & ; x_5 < 75,74 \\ 67,472x_5 - 5230,68 & ; 75,74 \leq x_5 < 78,64 \\ -105,658x_5 + 8384,263 & ; 78,64 \leq x_5 < 81,55 \\ 352,525x_5 - 28980,56 & ; x_5 \geq 81,55 \end{cases}$$

Pada interval pertama yaitu persentase pelaku kecelakaan tidak memiliki SIM yang kurang dari 75,74 persen, apabila terjadi kenaikan pada persentase pelaku kecelakaan tidak memiliki SIM sebesar satu persen, maka akan terjadi penurunan angka kecelakaan lalu lintas di Jawa Timur sebesar 1,589 kejadian kecelakaan per 100 kilometer. Terdapat 35 kabupaten/kota di Jawa Timur yang masuk dalam interval ini yaitu Kabupaten Sidoarjo, Kabupaten Bojonegoro, Kabupaten Ngawi, Kabupaten Sumenep, Kabupaten Pasuruan, Kabupaten Situbondo, Kota Malang, Kabupaten Probolinggo, Kabupaten Bangkalan, Kabupaten Malang, Kota Probolinggo, Kabupaten Pamekasan, Kota Mojokerto, Kabupaten Banyuwangi, Kabupaten Blitar, Kabupaten Lumajang, Kota Kediri, Kabupaten Pacitan, Kabupaten Nganjuk, Kabupaten Kediri, Kabupaten Trenggalek, Kabupaten Tulungagung, Kota Blitar, Kota Batu, Kabupaten Lamongan, Kabupaten Magetan, Kabupaten Ponorogo, Kabupaten Bondowoso, Kota Pasuruan, Kabupaten Tuban, Kabupaten Mojokerto, Kota Madiun, Kabupaten Madiun, Kabupaten Gresik, dan Kota Surabaya.

Untuk interval ketiga yaitu persentase pelaku kecelakaan tidak memiliki SIM yang berada pada interval 78,64 hingga 81,55 persen, apabila terjadi kenaikan sebesar satu persen, maka akan terjadi penurunan angka kecelakaan lalu lintas sebesar 105,658 kejadian kecelakaan per 100 kilometer jalan.

Terdapat 2 kabupaten/kota yang masuk dalam interval ini yaitu Kabupaten Jombang dan Kabupaten Sampang.

Selanjutnya, dengan persentase pelaku kecelakaan tidak memiliki SIM yang berada lebih dari 81,55 persen, apabila terjadi kenaikan sebesar satu persen maka akan terjadi kenaikan angka kecelakaan lalu lintas di Jawa Timur sebesar 352,525 kejadian kecelakaan per 100 kilometer jalan. Hanya terdapat satu kabupaten yang masuk dalam interval ini yaitu Kabupaten Jember. Pengelompokan kabupaten/kota di Jawa Timur berdasarkan variabel persentase pelaku kecelakaan tidak memiliki SIM yang mempengaruhi angka kecelakaan lalu lintas di Jawa Timur secara visual disajikan pada Gambar 4.18 sebagai berikut.



Gambar 4.18 Persebaran Kab./Kota Berdasarkan Pelaku Tanpa SIM

BAB V

KESIMPULAN DAN SARAN

5.1 Kesimpulan

Berdasarkan hasil analisis dan pembahasan yang telah dilakukan, dapat disimpulkan beberapa hal sebagai berikut.

1. Pada tahun 2015, angka kecelakaan lalu lintas terendah di Provinsi Jawa Timur berada di Kabupaten Blitar, sedangkan angka kecelakaan tertinggi terdapat di Kota Pasuruan. Terdapat 13 kabupaten/kota di Jawa Timur yang memiliki angka kecelakaan lalu lintas lebih tinggi dari rata-rata dan 25 kabupaten/kota di Jawa Timur yang memiliki angka kecelakaan lalu lintas lebih rendah dari rata-rata. Kepadatan penduduk yang paling rendah terdapat di Kabupaten Pacitan, sedangkan kepadatan penduduk yang paling tinggi terdapat di Kota Surabaya. Persentase pelaku usia remaja tertinggi sebesar terdapat di Kabupaten Tulungagung, sedangkan Kabupaten Sumenep merupakan yang terendah. Persentase terendah pada variabel pelaku kecelakaan berpendidikan terakhir di atas SMP adalah Kabupaten Sampang, sedangkan persentase tertinggi terdapat di Kabupaten Sidoarjo. Rasio jenis kelamin pelaku terendah adalah Kabupaten Jember dan yang tertinggi merupakan Kabupaten Madiun, dimana pelaku kecelakaan laki-laki lebih dominan daripada perempuan. Nilai terendah pada persentase pelaku yang tidak memiliki SIM ada di Kabupaten Sidoarjo, sedangkan Kabupaten Jember merupakan kabupaten dengan persentase tertinggi.
2. Model regresi nonparametrik Spline terbaik dalam pemodelan angka kecelakaan lalu lintas di Provinsi Jawa Timur adalah dengan menggunakan kombinasi titik knot (1,3,3,3,3). Model regresi nonparametrik Spline yang diperoleh adalah sebagai berikut.

$$\begin{aligned}\hat{y} = & 93,015 + 0,029x_1 - 0,488(x_1 - 8102,86)_+^1 + 3,566x_2 + 316,341(x_2 - 23,26)_+^1 - \\ & 532,815(x_2 - 24,23)_+^1 - 211,674(x_2 - 25,19)_+^1 + 0,177x_3 - 118,672(x_3 - 90,01)_+^1 + \\ & 338,428(x_3 - 93,36)_+^1 - 463,59(x_3 - 96,71)_+^1 - 0,41x_4 + 68,385(x_4 - 36,04)_+^1 + \\ & 41,031(x_4 - 37,35)_+^1 + 13,677(x_4 - 38,66)_+^1 - 1,589x_5 + 69,061(x_5 - 75,74)_+^1 - \\ & 173,13(x_5 - 78,64)_+^1 + 458,183(x_5 - 81,55)_+^1\end{aligned}$$

dengan keterangan

x_1 : Kepadatan penduduk

x_2 : Persentase pelaku usia remaja

x_3 : Persentase pendidikan terakhir pelaku lebih dari SMP

x_4 : Rasio jenis kelamin pelaku

x_5 : Persentase pelaku kecelakaan tidak memiliki SIM

Kelima variabel prediktor berpengaruh terhadap angka kecelakaan lalu lintas di Provinsi Jawa Timur yaitu variabel kepadatan penduduk, persentase pelaku kecelakaan usia remaja, persentase pelaku kecelakaan berpendidikan terakhir lebih dari SMP, rasio jenis kelamin pelaku, dan persentase pelaku yang tidak memiliki SIM dengan nilai koefisien determinasi yang dihasilkan dari model ini adalah sebesar 86,73 persen.

5.2 Saran

Saran yang dapat direkomendasikan terhadap pembaca dan peneliti selanjutnya adalah sebagai berikut.

1. Penelitian selanjutnya diharapkan menambah faktor-faktor lain yang berkaitan, sehingga dapat lebih menggambarkan kondisi yang mempengaruhi angka kecelakaan lalu lintas di Provinsi Jawa Timur tiap kabupaten/kota.
2. Upaya yang perlu dilakukan pemerintah daerah Provinsi Jawa Timur, salah satunya dengan mengadakan sosialisasi yang berkaitan dengan pemahaman dasar tentang berlalu-lintas dengan tertib, mengingat angka kecelakaan lalu lintas yang terjadi di Provinsi Jawa Timur sangat tinggi.

DAFTAR PUSTAKA

- Anwar, D.R. (2017). *Analisis Risiko Kecelakaan Lalu Lintas Berdasar Tipe Kepribadian dan Pelanggaran pada Pengendara Sepeda Motor*. Surabaya: Universitas Airlangga..
- Badan Pusat Statistik. (2013). *Publikasi Statistik Transportasi 2013*. Diakses melalui <http://www.bps.go.id> download tanggal 28 Februari 2018.
- _____. (2014). *Publikasi Statistik Transportasi 2014*. Diakses melalui <http://www.bps.go.id> download tanggal 28 Februari 2018.
- _____. (2015). *Kependudukan: Proyeksi Penduduk Provinsi Jawa Timur*. Diakses melalui <http://www.bps.go.id> download tanggal 28 Februari 2018.
- _____. (2015). *Publikasi Statistik Kesejahteraan Rakyat*. Diakses melalui <http://www.bps.go.id> download tanggal 28 Februari 2018.
- Budiantara, I. N., (2001). *Estimasi Parametrik dan Nonparametrik untuk Pendekatan Kurva Regresi*. Pembicara Utama Seminar Nasional Statistika V, Departemen Statistika, Fakultas Matematika, Komputasi, dan Sains Data, Institut Teknologi Sepuluh Nopember.
- _____. (2006). *Model Spline dengan Knot Optimal*. Jurnal Ilmu Dasar. FMIPA, Universitas Jember, Vol.7, Hal.77-85.
- _____. (2009). *Spline dalam Regresi Nonparametrik: Sebuah Pemodelan Statistika Masa Kini dan Masa Mendatang*. Pidato Pengukuhan Guru Besar Fakultas Matematika dan Ilmu Pengetahuan Alam, Institut Teknologi Sepuluh Nopember: Surabaya.
- Daniel, W.W. (1989). *Statistika Non Parametrik*. Jakarta: Gramedia.
- Dewi, P.L.A. (2016). *Pemodelan Faktor Penyebab Kecelakaan Lalu Lintas Berdasarkan Metode Geographically Weighted*

- Regression di Jawa Timur*. Surabaya: Institut Teknologi Sepuluh Nopember.
- Diana, Z.P. (2017). *Pemodelan Tingkat Keparahan Korban Kecelakaan Lalu Lintas di Kabupaten Pasuruan Dengan Pendekatan Regresi Logistik Ordinal*. Surabaya: Universitas Airlangga.
- Drapper, N. R. Dan Smith H. (1992). *Analisis Regresi Terapan Edisi Kedua*. Jakarta: PT. Gramedia Pustaka Utama.
- Eubank, R. L. (1988). *Spline Smoothing and Nonparametric Regression*. New York: Marcel Dekker.Inc.
- Goodwin, G.C., Schoby, J., & Council W. (2014). *A Hot Spot Analysis of Teenage Crashes: An Assessment of Crashes in Houston, Texas*. Texas: Texas Southern University.
- Gujarati, D. (2004). *Basic Economics fourth edition*. New York: McGraw-Hill.Inc.
- Gumawang, J. (2015). *Kajian Tingkat Kerawanan Kecelakaan Lalu Lintas Sebagian Ruas Jalan di Kota Tangerang Secara Spasial*. Yogyakarta: Universitas Gadjah Mada.
- Hardle, W. (1990). *Applied Nonparametric Regression*. New York: Cambridge University Press.
- Hobbs. (1995). *Perencanaan dan Teknik Lalu Lintas*. Yogyakarta: Gadjah Mada University Press.
- Kartika, M. (2009). *Analisis Faktor Penyebab Kecelakaan Lalu Lintas pada Pengendara Sepeda Motor di Wilayah Depok*. Depok: Universitas Indonesia.
- Wulandari, K. (2016). *Pemodelan Faktor-Faktor yang Mempengaruhi Angka Morbiditas di Provinsi Jawa Timur Menggunakan Regresi Nonparametrik Spline*. Surabaya: Institut Teknologi Sepuluh Nopember.
- Marta. (2017). *Data Indonesia dan Dunia, Korban Laka Lantas Usia Remaja Urutan Kedua Terbanyak*. Diakses melalui (<http://www.beraunews.com/lipsus-real-life/liputan-khusus/3002-data-indonesia-dan-dunia-korban-laka-lantas-usia-remaja-urutan-kedua-terbanyak>) [15 April 2018].

- Paranoan, N. R. (2017). *Pemodelan Faktor-Faktor yang Memengaruhi IPM di Provinsi Sumatera Utara Menggunakan Regresi Nonparametrik Spline*. Surabaya: Institut Teknologi Sepuluh Nopember.
- Rahmadi, S. (2011). *Kecelakaan Lalu Lintas di Kecamatan Duren Sawit Jakarta Timur*. Depok: Universitas Indonesia
- Rahmayanti, W., (2014). *Pemodelan Faktor Penyebab Kecelakaan Lalu Lintas Dengan Metode Regresi Logistik Multinomial*. Surabaya: Institut Teknologi Sepuluh Nopember.
- Satria, G. (2014). *Tingginya Angka Laka Tidak Punya SIM*. Diakses (<https://autotekno.sindonews.com/read/tingginya-angka-kecelakaan-karena-tidak-punya-sim>) [16 April 2018]
- Triana, E.P.Y., Kusdarwati,H., dan Pramoedyo,H. (2013). *Pemodelan Jumlah Kecelakaan Pengendara Sepeda Motor di Daerah Black Spot Kota Malang Menggunakan Generalized Poisson Regression*. Malang: Universitas Brawijaya.
- Wahba, G. (1990). *Spline Models For Observational Data*. SIAM Philadelphia. CBMS-NSF Regional Conference Series in Applied Mathematics, Vol. 59.
- Walpole, R. E. (1995). *Pengantar Statistika*. Jakarta: Gramedia Pustaka Utama.
- Wei, W. W. (2006). *Time Series Analysis: Univariate and Multivariate Methods (2nd ed.)*. Pearson Addison Wesley.

(Halaman ini sengaja dikosongkan)

LAMPIRAN

Lampiran 1. Data Angka Kecelakaan Lalu Lintas di Provinsi Jawa Timur dengan Faktor - Faktor yang Mempengaruhi Tahun 2015

Kab/kota	y	x_1	x_2	x_3	x_4	x_5
Pacitan	23.058	388	9.444	84.127	22.49	66.111
Ponorogo	57.853	613	14.463	82.828	15.318	63.115
Trenggalek	37.907	554	25.588	80.580	12.032	54.142
Tulungagung	39.259	887	25.676	89.503	8.708	53.763
Blitar	6.471	653	7.059	81.148	16.379	62.548
Kediri	41.093	1016	12.212	79.550	8.85	57.515
Malang	32.479	736	4.835	57.175	10.263	62.664
Lumajang	35.076	571	7.267	46.334	14.388	67.908
Jember	32.310	726	4.839	93.381	7.224	83.005
Banyuwangi	24.751	444	8.760	74.791	7.739	63.857
Bondowoso	28.263	485	5.279	59.524	7.463	70.262
Situbondo	43.202	406	2.899	89.021	8.602	66.265
Probolinggo	55.993	668	4.513	54.327	9.163	66.274
Pasuruan	35.105	1064	2.881	70.081	13.289	49.591
Sidoarjo	120.424	2945	7.853	98.390	14.286	11.782
Mojokerto	68.048	1109	19.532	90.738	11.355	51.857
Jombang	161.045	1114	6.858	70.374	18.182	79.307
Nganjuk	68.155	811	11.910	60.107	18.82	71.502
Madiun	44.277	604	5.729	92.742	39.316	48.462
Magetan	94.140	889	7.742	72.769	12.453	55.745
Ngawi	102.416	595	13.958	69.659	16.404	32.212
Bojonegoro	108.462	534	7.776	73.323	28.799	22.615
Tuban	146.670	583	23.634	90.588	20.513	68.861
Lamongan	173.438	675	6.452	83.810	21.106	57.726
Gresik	98.993	1015	4.587	95.550	19.064	62.700
Bangkalan	29.250	732	4.813	40.513	26.768	52.632
Sampang	15.957	759	4.348	16.304	10.577	80.435
Pamekasan	41.182	1062	6.373	69.231	15.457	71.220
Sumenep	7.445	514	1.980	59.434	9.04	69.608
Kota Kediri	144.960	4058	8.595	89.912	11.659	55.975
Kota Blitar	74.631	4179	7.143	77.174	20.354	74.725
Kota Malang	232.988	7739	4.444	83.784	12.547	49.265
Kota Probolinggo	145.376	4241	4.895	52.431	13.828	60.764
Kota Pasuruan	439.602	5127	3.691	97.241	12.018	44.518
Kota Mojokerto	273.585	6285	6.590	74.177	10.87	61.648
Kota Madiun	61.077	5147	5.118	91.339	7.91	42.913
Kota Surabaya	62.352	8606	3.693	79.292	13.924	37.685
Kota Batu	43.425	993	5.294	80.925	9.574	50.867

Keterangan:

y : Angka kecelakaan lalu lintas

x_1 : Kepadatan penduduk

x_2 : Persentase pelaku usia remaja

x_3 : Persentase pendidikan terakhir pelaku lebih dari SMP

x_4 : Rasio jenis kelamin pelaku kecelakaan

x_5 : Persentase pelaku kecelakaan tidak memiliki SIM

Lampiran 2. Program Pemilihan Titik Knot Optimal dengan Satu Titik Knot Menggunakan *Software R*

```

gcv1=function(para)
{
data=read.table("e://data.txt",header=FALSE)
data=as.matrix(data)
p=length(data[,1])
q=length(data[1,])
m=ncol(data)-para-1
dataA=data[, (para+2):q]
F=matrix(0,nrow=p,ncol=p)
diag(F)=1
nk= length(seq(min(data[,2]),max(data[,2]),length.out=50))
knot1=matrix(ncol=m,nrow=nk)
for (i in (1:m))
{
for (j in (1:nk))
{
a=seq(min(dataA[,i]),max(dataA[,i]),length.out=50)
knot1[j,i]=a[j]
}
}
a1=length(knot1[,1])
knot1=knot1[2:(a1-1),]
aa=rep(1,p)
data1=matrix(ncol=m,nrow=p)
data2=data[,2:q]
a2=nrow(knot1)
GCV=rep(NA,a2)
Rsq=rep(NA,a2)
for (i in 1:a2)
{
for (j in 1:m)
{

```

```

for (k in 1:p)
{
if(data[k,(j+para+1)]<knot1[i,j])data1[k,j]=0
else data1[k,j]=data[k,(j+para+1)]-knot1[i,j]
}
}
mx=cbind(aa,data2,data1)
mx=as.matrix(mx)
C=pinv(t(mx)%*%mx)
B=C%*%(t(mx)%*%data[,1])
yhat=mx%*%B
SSE=0
SSR=0
for (r in (1:p))
{
sum=(data[r,1]-yhat[r,])^2
sum1=(yhat[r,]-mean(data[,1]))^2
SSE=SSE+sum
SSR=SSR+sum1
}
Rsqr[i]=(SSR/(SSE+SSR))*100
MSE=SSE/p
A=mx%*%C%*%t(mx)
A1=(F-A)
A2=(sum(diag(A1))/p)^2
GCV[i]=MSE/A2
}
GCV=as.matrix(GCV)
Rsqr=as.matrix(Rsqr)
cat("=====","\n")
cat("Nilai Knot dengan Spline linear 1 knot","\n")
cat("=====","\n")
print (knot1)
cat("=====","\n")
cat("Rsqr dengan Spline linear 1 knot","\n")

```

```

cat("=====", "\n")
print (Rsqr)
cat("=====", "\n")
cat("HASIL GCV dengan Spline linear 1 knot", "\n")
cat("=====", "\n")
print (GCV)
s1=min(GCV)
print(max(Rsq))
cat("=====", "\n")
cat("HASIL GCV terkecil dengan Spline linear 1 knot", "\n")
cat("=====", "\n")
cat(" GCV =", s1, "\n")
write.csv(GCV, file="e:/coba/output GCV1.csv")
write.csv(Rsq, file="e:/coba/output Rsq1.csv")
write.csv(knot1, file="e:/coba/output knot1.csv")
}

```

Lampiran 3. Program Pemilihan Titik Knot Optimal dengan Dua Titik Knot Menggunakan *Software R*

```

gcv2=function(para)
{
data=read.table("e://data.txt",header=FALSE)
data=as.matrix(data)
p=length(data[,1])
q=length(data[1,])
m=ncol(data)-para-1
F=matrix(0,nrow=p,ncol=p)
dataA=data[, (para+2):q]
diag(F)=1
nk= length(seq(min(data[,2]),max(data[,2]),length.out=50))
knot=matrix(ncol=m,nrow=nk)
for (i in (1:m))
{
for (j in (1:nk))
{
a=seq(min(dataA[,i]),max(dataA[,i]),length.out=50)
knot[j,i]=a[j]
}
}
z=(nk*(nk-1)/2)
knot2=cbind(rep(NA,(z+1)))
for (i in (1:m))
{
knot1=rbind(rep(NA,2))
for ( j in 1:(nk-1))
{
for (k in (j+1):nk)
{
xx=cbind(knot[j,i],knot[k,i])
knot1=rbind(knot1,xx)
}
}
knot2=cbind(knot2,knot1)
}
knot2=knot2[2:(z+1),2:(2*m+1)]

```



```

aa=rep(1,p)
data2=matrix(ncol=(2*m),nrow=p)
data1=data[,2:q]
a1=length(knot2[,1])
GCV=rep(NA,a1)
Rsq=rep(NA,a1)
for (i in 1:a1)
{
  for (j in 1:(2*m))
  {
    if(mod(j,2)==1) b=floor(j/2)+1 else b=j/2
    for (k in 1:p)
    {
      if (data1[k,b]<knot2[i,j]) data2[k,j]=0
      else data2[k,j]=data1[k,b]-knot2[i,j]
    }
  }
  mx=cbind(aa,data1,data2)
  mx=as.matrix(mx)
  C=pinv(t(mx)%*%mx)
  B=C%*%(t(mx)%*%data[,1])
  yhat=mx%*%B
  SSE=0
  SSR=0
  for (r in (1:p))
  {
    sum=(data[r,1]-yhat[r,])^2
    sum1=(yhat[r,]-mean(data[,1]))^2
    SSE=SSE+sum
    SSR=SSR+sum1
  }
  Rsq[i]=(SSR/(SSE+SSR))*100
  MSE=SSE/p
  A=mx%*%C%*%t(mx)
  A1=(F-A)
  A2=(sum(diag(A1))/p)^2
  GCV[i]=MSE/A2
}
GCV=as.matrix(GCV)

```

```

Rsq=as.matrix(Rsq)
cat("=====", "\n")
cat("Nilai Knot dengan Spline linear 2 knot", "\n")
cat("=====", "\n")
print (knot2)
cat("=====", "\n")
cat("Rsq dengan Spline linear 2 knot", "\n")
cat("=====", "\n")
print (Rsq)
cat("=====", "\n")
cat("HASIL GCV dengan Spline linear 2 knot", "\n")
cat("=====", "\n")
print (GCV)
s1=min(GCV)
cat("=====", "\n")
cat("HASIL GCV terkecil dengan Spline linear 2 knot", "\n")
cat("=====", "\n")
cat(" GCV =", s1, "\n")
write.csv(GCV, file="e:/coba/output GCV2.csv")
write.csv(Rsq, file="e:/coba/output Rsq2.csv")
write.csv(knot2, file="e:/coba/output knot2.csv")
}

```

Lampiran 4. Program Pemilihan Titik Knot Optimal dengan Tiga Titik Knot Menggunakan *Software R*

```

gcv3=function(para)
{
data=read.table("e://data.txt",header=FALSE)
data=as.matrix(data)
p=length(data[,1])
q=length(data[,2])
m=ncol(data)-para-1
F=matrix(0,nrow=p,ncol=p)
dataA=data[, (para+2):q]
diag(F)=1
nk= length(seq(min(data[,2]),max(data[,2]),length.out=50))
knot=matrix(ncol=m,nrow=nk)
for (i in (1:m))
{
for (j in (1:nk))
{
a=seq(min(dataA[,i]),max(dataA[,i]),length.out=50)
knot[j,i]=a[j]
}
}
knot=knot[2:(nk-1),]
a2=nrow(knot)
z=(a2*(a2-1)*(a2-2)/6)
knot1=cbind(rep(NA,(z+1)))
for (i in (1:m))
{
knot2=rbind(rep(NA,3))
for ( j in 1:(a2-2))
{
for (k in (j+1):(a2-1))
{
for (g in (k+1):a2)
{
xx=cbind(knot[j,i],knot[k,i],knot[g,i])
knot2=rbind(knot2,xx)
}
}
}
}
}

```

```

}
}
knot1=cbind(knot1,knot2)
}
knot1=knot1[2:(z+1),2:(3*m+1)]
aa=rep(1,p)
data1=matrix(ncol=(3*m),nrow=p)
data2=data[, (para+2):q]
a1=length(knot1[,1])
GCV=rep(NA,a1)
Rsqr=rep(NA,a1)
for (i in 1:a1)
{
  for (j in 1:ncol(knot1))
  {
    b=ceiling(j/3)
    for (k in 1:p)
    {
      if (data2[k,b]<knot1[i,j]) data1[k,j]=0
      else data1[k,j]=data2[k,b]-knot1[i,j]
    }
  }
  mx=cbind(aa,data[,2:q],data1)
  mx=as.matrix(mx)
  C=pinv(t(mx)%*%mx)
  B=C%*%(t(mx)%*%data[,1])
  yhat=mx%*%B
  SSE=0
  SSR=0
  for (r in (1:p))
  {
    sum=(data[r,1]-yhat[r,])^2
    sum1=(yhat[r,]-mean(data[,1]))^2
    SSE=SSE+sum
    SSR=SSR+sum1
  }
  Rsqr[i]=(SSR/(SSE+SSR))*100
  MSE=SSE/p
  A=mx%*%C%*%t(mx)

```

```

A1=(F-A)
A2=(sum(diag(A1))/p)^2
GCV[i]=MSE/A2
}
GCV=as.matrix(GCV)
Rsqr=as.matrix(Rsq)
cat("=====", "\n")
cat("Nilai Knot dengan Spline linear 3 knot", "\n")
cat("=====", "\n")
print (knot1)
cat("=====", "\n")
cat("Rsqr dengan Spline linear 3 knot", "\n")
cat("=====", "\n")
print (Rsqr)
r=max(Rsqr)
print (r)
cat("=====", "\n")
cat("HASIL GCV dengan Spline linear 3 knot", "\n")
cat("=====", "\n")
print (GCV)
s1=min(GCV)
cat("=====", "\n")
cat("HASIL GCV terkecil dengan Spline linear 3 knot", "\n")
cat("=====", "\n")
cat(" GCV =", s1, "\n")
write.csv(GCV, file="e:/coba/output GCV3.csv")
write.csv(Rsqr, file="e:/coba/output Rsqr3.csv")
write.csv(knot1, file="e:/coba/output knot3.csv")
}

```

Lampiran 5. Program Pemilihan Titik Knot Optimal dengan Kombinasi Titik Knot Menggunakan *Software R*

```

gcvkom=function(para)
{
data=read.table("e://data.txt",header=FALSE)
data=as.matrix(data)
p1=length(data[,1])
q1=length(data[1,])
v=para+2
F=matrix(0,nrow=p1,ncol=p1)
diag(F)=1
x1=read.table("e://x1.txt",header=FALSE)
x2=read.table("e://x2.txt",header=FALSE)
x3=read.table("e://x3.txt",header=FALSE)
x4=read.table("e://x4.txt",header=FALSE)
x5=read.table("e://x5.txt",header=FALSE)
n2=nrow(x1)
a=matrix(nrow=5,ncol=3^5)
m=0
for (i in 1:3)
for (j in 1:3)
for (k in 1:3)
for (l in 1:3)
for (z in 1:3)
{
m=m+1
a[,m]=c(i,j,k,l,z)
}
a=t(a)
GCV=matrix(nrow=nrow(x1),ncol=3^5)
for (i in 1:3^5)
{
for (h in 1:nrow(x1))
{
if (a[i,1]==1)
{
gab=as.matrix(x1[,1])
gen=as.matrix(data[,v])

```

```

aa=matrix(nrow=nrow(x1)*nrow(data),ncol=1)
for (j in 1:1)
for (w in 1:nrow(data))
{
if (gen[w,j]<gab[h,j]) aa[w,j]=0 else aa[w,j]=gen[w,j]-gab[h,j]
}
}
else
if (a[i,1]==2)
{
gab=as.matrix(x1[,2:3])
gen=as.matrix(cbind(data[,v],data[,v]))
aa=matrix(nrow=nrow(x1)*nrow(data),ncol=2)
for (j in 1:2)
for (w in 1:nrow(data))
{
if (gen[w,j]<gab[h,j]) aa[w,j]=0 else aa[w,j]=gen[w,j]-gab[h,j]
}
}
else
{
gab=as.matrix(x1[,4:6])
gen=as.matrix(cbind(data[,v],data[,v],data[,v]))
aa=matrix(nrow=nrow(x1)*nrow(data),ncol=3)
for (j in 1:3)
for (w in 1:nrow(data))
{
if (gen[w,j]<gab[h,j]) aa[w,j]=0 else aa[w,j]=gen[w,j]-gab[h,j]
}
}
if (a[i,2]==1)
{
gab=as.matrix(x2[,1] )
gen=as.matrix(data[, (v+1)])
bb=matrix(nrow=nrow(x1)*nrow(data),ncol=1)
for (j in 1:1)
for (w in 1:nrow(data))
{
if (gen[w,j]<gab[h,j]) bb[w,j]=0 else bb[w,j]=gen[w,j]-gab[h,j]
}
}
}

```

```

}
}
else
if (a[i,2]==2)
{
gab=as.matrix(x2[,2:3] )
gen=as.matrix(cbind(data[, (v+1)], data[, (v+1)]))
bb=matrix(nrow=nrow(x1)*nrow(data), ncol=2)
for (j in 1:2)
for (w in 1:nrow(data))
{
if (gen[w,j]<gab[h,j]) bb[w,j]=0 else bb[w,j]=gen[w,j]-gab[h,j]
}
}
else
{
gab=as.matrix(x2[,4:6])
gen=as.matrix(cbind(data[, (v+1)], data[, (v+1)], data[, (v+1)]))
bb=matrix(nrow=nrow(x1)*nrow(data), ncol=3)
for (j in 1:3)
for (w in 1:nrow(data))
{
if (gen[w,j]<gab[h,j]) bb[w,j]=0 else bb[w,j]=gen[w,j]-gab[h,j]
}
}
if (a[i,3]==1)
{
gab=as.matrix(x3[,1] )
gen=as.matrix(data[, (v+2)])
cc=matrix(nrow=nrow(x1)*nrow(data), ncol=1)
for (j in 1:1)
for (w in 1:nrow(data))
{
if (gen[w,j]<gab[h,j]) cc[w,j]=0 else cc[w,j]=gen[w,j]-gab[h,j]
}
}
else
if (a[i,3]==2)
{

```



```

gab=as.matrix(x3[,2:3] )
gen=as.matrix(cbind(data[, (v+2)], data[, (v+2)]))
cc=matrix(nrow=nrow(x1)*nrow(data), ncol=2)
for (j in 1:2)
for (w in 1:nrow(data))
{
if (gen[w,j]<gab[h,j]) cc[w,j]=0 else cc[w,j]=gen[w,j]-gab[h,j]
}
}
else
{
gab=as.matrix(x3[,4:6])
gen=as.matrix(cbind(data[, (v+2)], data[, (v+2)], data[, (v+2)]))
cc=matrix(nrow=nrow(x1)*nrow(data), ncol=3)
for (j in 1:3)
for (w in 1:nrow(data))
{
if (gen[w,j]<gab[h,j]) cc[w,j]=0 else cc[w,j]=gen[w,j]-gab[h,j]
}
}
if (a[i,4]==1)
{
gab=as.matrix(x4[,1] )
gen=as.matrix(data[, (v+3)])
dd=matrix(nrow=nrow(x1)*nrow(data), ncol=1)
for (j in 1:1)
for (w in 1:nrow(data))
{
if (gen[w,j]<gab[h,j]) dd[w,j]=0 else dd[w,j]=gen[w,j]-gab[h,j]
}
}
else
if (a[i,4]==2)
{
gab=as.matrix(x4[,2:3] )
gen=as.matrix(cbind(data[, (v+3)], data[, (v+3)]))
dd=matrix(nrow=nrow(x1)*nrow(data), ncol=2)
for (j in 1:2)
for (w in 1:nrow(data))

```

```

{
if (gen[w,j]<gab[h,j]) dd[w,j]=0 else dd[w,j]=gen[w,j]-gab[h,j]
}
}
else
{
gab=as.matrix(x4[,4:6])
gen=as.matrix(cbind(data[, (v+3)], data[, (v+3)], data[, (v+3)]))
dd=matrix(nrow=nrow(x1)*nrow(data), ncol=3)
for (j in 1:3)
for (w in 1:nrow(data))
{
if (gen[w,j]<gab[h,j]) dd[w,j]=0 else dd[w,j]=gen[w,j]-gab[h,j]
}
}
if (a[i,5]==1)
{
gab=as.matrix(x5[,1] )
gen=as.matrix(data[, (v+4)])
ee=matrix(nrow=nrow(x1)*nrow(data), ncol=1)
for (j in 1:1)
for (w in 1:nrow(data))
{
if (gen[w,j]<gab[h,j]) ee[w,j]=0 else ee[w,j]=gen[w,j]-gab[h,j]
}
}
else
if (a[i,5]==2)
{
gab=as.matrix(x5[,2:3] )
gen=as.matrix(cbind(data[, (v+4)], data[, (v+4)]))
ee=matrix(nrow=nrow(x1)*nrow(data), ncol=2)
for (j in 1:2)
for (w in 1:nrow(data))
{
if (gen[w,j]<gab[h,j]) ee[w,j]=0 else ee[w,j]=gen[w,j]-gab[h,j]
}
}
}
else

```

```

{
gab=as.matrix(x5[,4:6])
gen=as.matrix(cbind(data[, (v+4)], data[, (v+4)], data[, (v+4)]))
ee=matrix(nrow=nrow(x1)*nrow(data), ncol=3)
for (j in 1:3)
for (w in 1:nrow(data))
{
if (gen[w,j]<gab[h,j]) ee[w,j]=0 else ee[w,j]=gen[w,j]-gab[h,j]
}
}

ma=as.matrix(cbind(aa,bb,cc,dd,ee))
mx=cbind(rep(1,nrow(data)),data[,2:q1],na.omit(ma))
mx=as.matrix(mx)
C=pinv(t(mx)%*%mx)
B=C%*%(t(mx)%*%data[,1])
yhat=mx%*%B
SSE=0
SSR=0
for (r in 1:nrow(data))
{
sum=(data[r,1]-yhat[r,])^2
sum1=(yhat[r,]-mean(data[,1]))^2
SSE=SSE+sum
SSR=SSR+sum1
}
Rsqr=(SSR/(SSE+SSR))*100
MSE=SSE/p1
A=mx%*%C%*%t(mx)
A1=(F-A)
A2=(sum(diag(A1))/p1)^2
GCV[h,i]=MSE/A2
}

if (a[i,1]==1) sp=x1[,1] else
if (a[i,1]==2) sp=x1[,2:3] else
sp=x1[,4:6]
if (a[i,2]==1) spl=x2[,1] else
if (a[i,2]==2) spl=x2[,2:3] else

```

```

spl=x2[,4:6]
if (a[i,3]==1) splin=x3[,1] else
if (a[i,3]==2) splin=x3[,2:3] else
splin=x3[,4:6]
if (a[i,4]==1) spline=x4[,1] else
if (a[i,4]==2) spline=x4[,2:3] else
spline=x4[,4:6]
if (a[i,5]==1) splines=x5[,1] else
if (a[i,5]==2) splines=x5[,2:3] else
splines=x5[,4:6]
kkk=cbind(sp,spl,splin,spline,splines)
cat("=====")
", "\n")
print(i)
print(kkk)
print(Rsq)
}
write.csv(GCV,file="e:/coba/output GCV kombinasi.csv")
}

```

Lampiran 6. Program Pengujian Parameter Menggunakan *Software R*

```

uji=function(alpha,para)
{
data=read.table("e://data.txt")
knot=read.table("e://knot.txt")
data=as.matrix(data)
knot=as.matrix(knot)
ybar=mean(data[,1])
m=para+2
p=nrow(data)
q=ncol(data)
dataA=cbind(data[,m],data[,m+1],data[,m+1],data[,m+1],data[,m
+2],data[,m+2],data[,m+2],data[,m+3],data[,m+3],data[,m+3],dat
a[,m+4],data[,m+4],data[,m+4])
dataA=as.matrix(dataA)
satu=rep(1,p)
n1=ncol(knot)
data.knot=matrix(ncol=n1,nrow=p)
for (i in 1:n1)
{
for(j in 1:p)
{
if(dataA[j,i]<knot[1,i]) data.knot[j,i]=0
else data.knot[j,i]=dataA[j,i]-knot[1,i]
}
}
mx=cbind(satu,data[,2],data.knot[,1],data[,3],data.knot[,2:4],data[
,4],data.knot[,5:7],data[,5],data.knot[,8:10],data[,6],data.knot[,11:
13])
mx=as.matrix(mx)
B=(pinv(t(mx)%*%mx))%*%t(mx)%*%data[,1]
cat("=====", "\n")
cat("Estimasi Parameter", "\n")
cat("=====", "\n")

```

```

print (B)
n1=nrow(B)
yhat=mx%*%B
res=data[,1]-yhat
SSE=sum((data[,1]-yhat)^2)
SSR=sum((yhat-ybar)^2)
SST=SSR+SSE
MSE=SSE/(p-n1)
MSR=SSR/(n1-1)
Rsqr=(SSR/(SSR+SSE))*100

#uji F (uji serentak)
Fhit=MSR/MSE
pvalue=pf(Fhit,(n1-1),(p-n1),lower.tail=FALSE)
if (pvalue<=alpha)
{
cat("-----", "\n")
cat("Kesimpulan hasil uji serentak", "\n")
cat("-----", "\n")
cat("Tolak Ho yakni minimal terdapat 1 prediktor yang
signifikan", "\n")
cat("", "\n")
}
else
{
cat("-----", "\n")
cat("Kesimpulan hasil uji serentak", "\n")
cat("-----", "\n")
cat("Gagal Tolak Ho yakni semua prediktor tidak berpengaruh
signifikan", "\n")
cat("", "\n")
}

#uji t (uji individu)
thit=rep(NA,n1)

```

```

pval=rep(NA,n1)
SE=sqrt(diag(MSE*(pinv(t(mx)%*%mx))))
cat("-----", "\n")
cat("Kesimpulan hasil uji individu", "\n")
cat("-----", "\n")
thit=rep(NA,n1)
pval=rep(NA,n1)
for (i in 1:n1)
{
  thit[i]=B[i,1]/SE[i]
  pval[i]=2*(pt(abs(thit[i]),(p-n1),lower.tail=FALSE))
  if (pval[i]<=alpha) cat("Tolak Ho yakni prediktor signifikan
dengan pvalue",pval[i],"\n") else cat("Gagal tolak Ho yakni
prediktor tidak signifikan dengan pvalue",pval[i],"\n")
}
thit=as.matrix(thit)
cat("=====","\n")
cat("nilai t hitung", "\n")
cat("=====","\n")
print (thit)
cat("-----", "\n")
cat("Analysis of Variance", "\n")
cat("=====","\n")
      cat("Sumber      df      SS      MS      Fhit", "\n")
      cat("Regresi      ",(n1-1)," ",SSR," ",MSR,"",Fhit," \n")
      cat("Error        ",p-n1," ",SSE,"",MSE," \n")
      cat("Total        ",p-1," ",SST," \n")
cat("=====","\n")
      cat("s=",sqrt(MSE),"      Rsq=",Rsq," \n")
      cat("pvalue(F)=",pvalue," \n")
write.csv(res,file="e:/coba/output uji residual.csv")
write.csv(pval,file="e:/coba/output uji pvalue.csv")
write.csv(mx,file="e:/coba/output uji mx.csv")
write.csv(yhat,file="e:/coba/output uji yhat.csv")
}

```

Lampiran 7. Uji Glejser Menggunakan *Software* Minitab

Regression Analysis: abs resi versus Kepadatan pe, Pelaku usia , ...

The regression equation is

abs resi = - 14.8 + 0.00077 Kepadatan penduduk (X1)
 - 0.384 Pelaku usia remaja (X2) + 0.345 pelaku
 di atas SMP (X3) + 0.123 rasio JK (X4) + 0.163
 pelaku tanpa sim (X5)

Predictor	Coef	SE Coef	T	P
Constant	-14.78	32.18	-0.46	0.649
Kepadatan penduduk (X1)	0.000766	0.002072	0.37	0.714
Pelaku usia remaja (X2)	-0.3841	0.7137	-0.54	0.594
pelaku di atas SMP (X3)	0.3446	0.2515	1.37	0.180
rasio JK (X4)	0.1226	0.6146	0.20	0.843
pelaku tanpa sim (X5)	0.1627	0.2983	0.55	0.589

S = 23.8778 R-Sq = 7.1% R-Sq(adj) = 0.0%

Analysis of Variance

Source	DF	SS	MS	F	P
Regression	5	1396.1	279.2	0.49	0.781
Residual Error	32	18244.8	570.2		
Total	37	19640.9			

Source	DF	Seq SS
Kepadatan penduduk (X1)	1	264.3
Pelaku usia remaja (X2)	1	16.4
pelaku di atas SMP (X3)	1	943.4
rasio JK (X4)	1	2.3
pelaku tanpa sim (X5)	1	169.6

Unusual Observations

Obs	Kepadatan penduduk (X1)	abs resi	Fit	SE Fit	Residual	St
24	675	123.10	24.12	6.58	98.98	
4.31R						

R denotes an observation with a large standardized residual.

Lampiran 8. Output Nilai GCV dengan Satu Titik Knot

No	GCV	x_1	x_2	x_3	x_4	x_5
1	8032,605	555,714	2,464	17,979	7,879	13,236
2	8017,297	723,429	2,947	19,654	8,534	14,689
3	7839,713	891,143	3,431	21,330	9,189	16,143
4	7853,898	1058,857	3,914	23,005	9,844	17,596
5	7781,409	1226,571	4,398	24,680	10,499	19,050
6	7779,556	1394,286	4,882	26,355	11,154	20,503
7	7800,988	1562,000	5,365	28,031	11,809	21,957
8	7945,410	1729,714	5,849	29,706	12,464	23,410
9	8133,215	1897,429	6,332	31,381	13,118	24,864
10	8244,480	2065,143	6,816	33,056	13,773	26,317
⋮	⋮	⋮	⋮	⋮	⋮	⋮
40	6249,899	7096,571	21,324	83,313	33,422	69,923
41	6131,908	7264,286	21,807	84,988	34,076	71,377
42	5977,798	7432,000	22,291	86,663	34,731	72,830
43	5809,157	7599,714	22,774	88,339	35,386	74,284
44	5626,401	7767,429	23,258	90,014	36,041	75,737
45	5359,521	7935,143	23,742	91,689	36,696	77,191
46	5329,136	8102,857	24,225	93,364	37,351	78,644
47	5496,004	8270,571	24,709	95,040	38,006	80,098
48	6258,352	8438,286	25,192	96,715	38,661	81,551

Lampiran 9. Output Nilai GCV dengan Dua Titik Knot

No	GCV	x_1	x_2	x_3	x_4	x_5
1	8518,701	388	1,98	16,304	7,224	11,782
		555,714	2,464	17,979	7,879	13,236
2	8401,389	388	1,98	16,304	7,224	11,782
		723,429	2,947	19,654	8,534	14,689
3	7989,853	388	1,98	16,304	7,224	11,782
		891,143	3,431	21,330	9,189	16,143
⋮	⋮	⋮	⋮	⋮	⋮	⋮
598	11539,074	2736	8,750	39,757	16,393	32,131
		3239,143	10,201	44,783	18,358	36,492
599	11705,803	2736	8,750	39,757	16,393	32,131
		3406,857	10,685	46,458	19,013	37,946
600	11892,871	2736	8,750	39,757	16,393	32,131
		3574,571	11,168	48,133	19,668	39,399
⋮	⋮	⋮	⋮	⋮	⋮	⋮
1223	3791,552	8270,571	24,709	95,040	38,006	80,098
		8438,286	25,192	96,715	38,661	81,551
1224	5496,004	8270,571	24,709	95,040	38,006	80,098
		8606,000	25,676	98,390	39,316	83,005
1225	6258,352	8438,286	25,192	96,715	38,661	81,551
		8606,000	25,676	98,390	39,316	83,005

Lampiran 10. Output Nilai GCV dengan Tiga Titik Knot

No	GCV	x_1	x_2	x_3	x_4	x_5
1	10553,510	555,714	2,464	17,979	7,879	13,236
		723,429	2,947	19,654	8,534	14,689
		891,143	3,431	21,330	9,189	16,143
2	11463,512	555,714	2,464	17,979	7,879	13,236
		723,429	2,947	19,654	8,534	14,689
		1058,857	3,914	23,005	9,844	17,596
3	11261,065	555,714	2,464	17,979	7,879	13,236
		723,429	2,947	19,654	8,534	14,689
		1226,571	4,398	24,680	10,499	19,050
⋮	⋮	⋮	⋮	⋮	⋮	⋮
1224	5496,004	7935,143	23,742	91,689	36,696	77,191
		8270,571	24,709	95,040	38,006	80,098
		8438,286	25,192	96,715	38,661	81,551
1225	6258,352	8102,857	24,225	93,364	37,351	78,644
		8270,571	24,709	95,040	38,006	80,098
		8438,286	25,192	96,715	38,661	81,551
⋮	⋮	⋮	⋮	⋮	⋮	⋮
8974	14189,396	2232,857	7,300	34,731	14,428	27,771
		2903,714	9,234	41,432	17,048	33,585
		4413,143	13,586	56,509	22,943	46,667
8975	13257,394	2232,857	7,300	34,731	14,428	27,771
		2903,714	9,234	41,432	17,048	33,585
		4580,857	14,070	58,185	23,597	48,120

Lampiran 11. Output Nilai GCV dengan Kombinasi Titik Knot (1, 3, 3, 3, 3)

No	GCV	Kombinasi	x_1	x_2	x_3	x_4	x_5
1	5329,136	11111	8102,857	24,225	93,364	37,351	78,644
2	5716,677	11112	8102,857	24,225	93,364	37,351	80,098
							81,551
3	5809,842	11113	8102,857	24,225	93,364	37,351	75,737
							78,644
							81,551
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
81	2794,335	13333	8102,857	23,258	90,014	36,041	75,737
				24,225	93,364	37,351	78,644
				25,192	96,715	38,661	81,551
82	5329,136	21111	8270,571	24,225	93,364	37,351	78,644
			8438,286				
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
242	3579,174	33332	7767,429	23,258	90,014	36,041	80,098
			8102,857	24,225	93,364	37,351	81,551
			8438,286	25,192	96,715	38,661	
243	2794,335	33333	7767,429	23,258	90,014	36,041	75,737
			8102,857	24,225	93,364	37,351	78,644
			8438,286	25,192	96,715	38,661	81,551

Lampiran 12. Output Estimasi Parameter dan Uji Signifikansi

Estimasi Parameter

```

=====
[1] 93.01462054
[2] 0.02942268
[3] -0.48793749
[4] 3.56561054
[5] 316.34136468
[6] -532.81519296
[7] -211.67351617
[8] 0.17744251
[9] -118.67240117
[10] 338.42752523
[11] -463.95923064
[12] -0.41026828
[13] 68.38493946
[14] 41.03108897
[15] 13.67702966
[16] -1.58940847
[17] 69.06068211
[18] -173.13040418
[19] 458.18313369
=====

```

Kesimpulan hasil uji serentak

Tolak H_0 yakni minimal terdapat 1 prediktor yang signifikan

Kesimpulan hasil uji individu

Gagal tolak H_0 yakni prediktor tidak signifikan dengan pvalue 0.2705383

Tolak H_0 yakni prediktor signifikan dengan pvalue 2.000495e-06

Tolak H_0 yakni prediktor signifikan dengan pvalue 0.0002277318

Gagal tolak H_0 yakni prediktor tidak signifikan dengan pvalue 0.1489198

Gagal tolak H_0 yakni prediktor tidak signifikan dengan pvalue 0.0546289
 Tolak H_0 yakni prediktor signifikan dengan pvalue 0.02652707
 Tolak H_0 yakni prediktor signifikan dengan pvalue 0.02345579
 Gagal tolak H_0 yakni prediktor tidak signifikan dengan pvalue 0.8040366
 Tolak H_0 yakni prediktor signifikan dengan pvalue 0.0005432058
 Tolak H_0 yakni prediktor signifikan dengan pvalue $5.05747e-05$
 Tolak H_0 yakni prediktor signifikan dengan pvalue $2.625388e-05$
 Gagal tolak H_0 yakni prediktor tidak signifikan dengan pvalue 0.7998753
 Tolak H_0 yakni prediktor signifikan dengan pvalue 0.005108487
 Tolak H_0 yakni prediktor signifikan dengan pvalue 0.005108487
 Tolak H_0 yakni prediktor signifikan dengan pvalue 0.005108487
 Gagal tolak H_0 yakni prediktor tidak signifikan dengan pvalue 0.05580103
 Tolak H_0 yakni prediktor signifikan dengan pvalue 0.02330448
 Gagal tolak H_0 yakni prediktor tidak signifikan dengan pvalue 0.07075879
 Tolak H_0 yakni prediktor signifikan dengan pvalue 0.01017978

=====

nilai t hitung

=====

[,1]
 [1,] 1.1348854
 [2,] 6.7232721
 [3,] -4.5325538
 [4,] 1.5044032
 [5,] 2.0480685
 [6,] -2.4050105
 [7,] -2.4638899
 [8,] 0.2516196
 [9,] -4.1506171
 [10,] 5.2034194
 [11,] -5.5016752
 [12,] -0.2570868
 [13,] 3.1641330
 [14,] 3.1641330

```
[15,] 3.1641330
[16,] -2.0372312
[17,] 2.4669737
[18,] -1.9143432
[19,] 2.8527962
```

 Analysis of Variance

```
=====
Sumber    df    SS    MS    Fhit
Regresi   18 232534.9 12918.61 6.89653
Error     19 35590.87 1873.204
Total     37 268125.8
=====
```

```
s= 43.28052 Rsq= 86.72605
pvalue(F)= 5.520303e-05
```

Lampiran 13. Output Residual

No	Residual	No	Residual
1	-15,6696	20	28,16199
2	-12,8644	21	-12,3063
3	-13,1698	22	6,758653
4	12,19825	23	3,222746
5	-39,1913	24	123,0956
6	-44,4281	25	-32,4247
7	-5,76644	26	-15,0162
8	4,964607	27	-0,0196
9	0,004936	28	1,451524
10	-21,1632	29	-3,95453
11	6,330285	30	-20,302
12	20,96024	31	-53,385
13	26,68826	32	-34,9913
14	-27,6508	33	3,074652
15	-9,68922	34	30,84232
16	29,67296	35	61,43279
17	0,015945	36	10,87431
18	19,51376	37	-2,48E-09
19	8,23E-05	38	-27,2659

Lampiran 14. Surat Keterangan Pengambilan Data

SURAT KETERANGAN

Saya yang bertanda tangan di bawah ini menerangkan bahwa :

1. Mahasiswa Statistika FMKSD-ITS dengan identitas berikut :

Nama : Nym Cista Striratna Dewi

NRP : 06211645000011

Telah mengambil data di instansi/perusahaan kami :

Nama Instansi : DITLANTAS POLDA JATIM

Divisi/ bagian : SUBDITBINGAKKUM

sejak bulan Januari 2018 sampai dengan Juni 2018 untuk keperluan

Tugas Akhir Semester Genap 2017/2018.

2. Tidak Keberatan/~~Keberatan~~* nama instansi dicantumkan dalam Tugas Akhir mahasiswa Statistika yang akan di simpan di Perpustakaan ITS dan dibaca di lingkungan ITS.
3. Tidak Keberatan/~~Keberatan~~* bahwa hasil analisis data dari instansi dipublikasikan dalam E journal ITS yaitu Jurnal Sains dan Seni ITS.

Surabaya, 25 Juli 2018

KASUBDITBINGAKKUM



M. BUDI HENDRAN, S.I.K. M.H.

*(coret yang tidak perlu)

NIK. AKBP NRP 70120957

(Halaman ini sengaja dikosongkan)

BIODATA PENULIS



Nym Cista Striratna Dewi dilahirkan di Pontianak, 7 Maret 1996. Anak ketiga dari I Ketut Murdana dan Endang Rini Susanti ini memiliki hobi berpetualang dan berenang. Penulis telah menempuh pendidikan formal di SD Negeri Wonokromo III Surabaya, SMP Negeri 32 Surabaya, dan SMA Negeri 16 Surabaya. Kemudian, penulis melanjutkan ke jenjang perguruan tinggi yaitu Diploma 3 Statistika ITS pada tahun 2013 melalui jalur SMITS, dan kemudian pada tahun 2016 penulis melanjutkan studi Strata 1 di jurusan dan almamater yang sama. Semasa kuliah, penulis mengikuti organisasi di luar lingkup jurusan yaitu sebagai staff ahli Departemen Umum TPKH-ITS selama dua periode kepengurusan yaitu tahun 2014-2015 dan tahun 2015-2016. Penulis pernah menjadi asisten dosen untuk mata kuliah statistika nonparametrik. Selain itu, penulis aktif dalam beberapa kepanitiaan di dalam kampus, pada tahun 2015 penulis menjadi *the head of instructor committee* GERIGI, kemudian menjadi panitia Pekan Raya Statistika ITS sebagai *public relation staff*, dan menjadi tim kreatif untuk *monthly magazine* HIMADATA - ITS. Apabila pembaca ingin berdiskusi maupun memberikan kritik dan saran mengenai tugas akhir ini, penulis dapat dihubungi melalui email: nymcista@gmail.com.

(Halaman ini sengaja dikosongkan)